Technological Change and Occupations over the Long Run*

Leonid Kogan† Dimitris Papanikolaou‡ Lawrence D. W. Schmidt§
Bryan Seegmiller¶

First Version: December 2019
Current Version: January 2021

Abstract

We construct occupation-specific indicators of technological change that span two centuries (1850-2010) using textual analysis of patent documents and occupation task descriptions. We find strong evidence that much of technical change has been displacive of labor during this period. Our occupation-level indices are significantly negatively correlated with future employment and wage growth; employment declines are especially pronounced among exposed occupations during innovation waves and around recessions. Aggregating at the industry level, our indices positively predict industry productivity and output but are associated with a decline in the labor share of output. Comparing the recent IT revolution to the Second Industrial Revolution (1880s) and the technology wave of the 1920s and 1930s, our measure uncovers an important difference: the first two waves consisted of innovations that were mainly related to occupations emphasizing (non-interpersonal) manual tasks, while the recent wave is composed of innovations that are significantly more related to cognitive tasks than the previous two waves.

*We are grateful to Daron Acemoglu and David Autor for valuable discussions and feedback and to Will Cong for generously sharing code.
†MIT Sloan School of Management and NBER
‡Kellogg School of Management and NBER
§MIT Sloan School of Management
¶MIT Sloan School of Management

Economists and workers alike have long worried about the employment prospects of occupations whose key functions can also be done by a machine, robot, or some other form of capital that substitutes for labor. These concerns may have been exacerbated due to recent advances in computers and software, yet they are not new. In 1930 John Maynard Keynes described this type of potential labor market risk when he said, "We are being afflicted with a new disease of technological unemployment...due to our discovery of means of economising the use of labor outrunning the pace at which we can find new uses for labor."[1] Yet, despite the importance of this question, evidence for technological displacement remains indirect. Our goal is to fill this gap by constructing time-series indices of the exposure of occupations to technological innovation over long periods of time.

Our starting point in constructing such a measure relies on estimating a distance metric between a given innovation (a patent) and an occupation based on its task description. To construct such a link, we exploit recent developments in natural language processing that allow us to compute the similarity between documents that can account for words that are similar, but not exactly, the same.[2] Armed with such a distance metric, we create time-series indices of the exposure of specific occupations to technological innovation by using the occupation-similarities to weigh 'breakthrough patents'—which are similarly defined using the patent text using the method of Kelly, Papanikolaou, Seru, and Taddy (2020).

In brief, our indices capture the extent to which specific occupations were exposed to important innovations in a given year. For instance, our index for "molders, shapers, and casters, except metal and plastic"—which occupation category prominently includes glass blowers as a sub-occupation—takes a relatively high value in the early 1900s because of similarity with patents such as US patent number 814,612, entitled "Method of Making glass sheets." This patent relates to a technology for making glass called the cylinder machine, which allowed glass manufacturers to replace the labor of skilled hand glass blowers in favor of a highly mechanized and capital-intensive production process. Jerome (1934) documents a dramatic transformation in the production process of the glass making industry as a result of the cylinder machine.[3] This example is illustrative of a more pervasive pattern that we

[1] See Keynes, The Economic Possibilities of our Grandchildren (1930).

[2] Standard methods for computing document similarity rely on measuring the amount of overlapping exact word matches, which cannot account for word similarities. Though this method has been used to some success in other domains, being able to account for words that are similar without having to exactly overlap is especially important given the very different structure and language in the text of patents and the occupation task descriptions found in the Dictionary of Occupational Titles. Our method uses pre-estimated vector representations of word meanings (called "word embeddings") and is straightforward to implement.

[3] Jerome (1934) writes: "By 1905 many hand plants had gone out of business, wages of blowers and gatherers were reduced 40 per cent, and the new machine may be said to have achieved commercial success . . . in the quarter century following the introduction of machine blowing, the window-glass industry, one of

uncover–our index of technological exposure based on patent and task textual similarity predicts employment declines dating back to the mid- to late-19th century and up through the early 21st century.

Examining our indicators, we find that, prior to 1980, innovation was consistently associated with manual physical tasks; by contrast, the innovations of the late 20th/early 21st century are much more related to cognitive tasks than before. This pattern is partly driven by the increased prevalence of breakthrough patents related to computers and electronics. Last, occupations that are associated with interpersonal tasks have consistently low exposures to innovation throughout the entire sample period.

Armed with measures of technology exposure, we next examine the relation between our indicators and labor market outcomes. We find that high values of our occupation-level indicators are followed by employment declines. This negative correlation is remarkable consistent over time —starting from the Second Industrial Revolution of the late 19th century to the present. We find some evidence that the negative relation between our measure and employment is stronger in recessions—consistent with the literature on job polarization.

Focusing on the more recent period for which wage earnings data is available (after 1980), we also document a negative relation between our technology exposure measure and wage growth. This predictability is complementary to the information contained in the routine-task intensity measure of Autor and Dorn (2013), the AI and robotics occupation exposure measure of Webb (2019) and is not driven by industry-specific trends. Further, the magnitudes are economically significant: controlling for differential industry trends, our point estimates imply that a one-standard deviation increase in our technology exposure predicts declines of about 0.8% in employment and 0.1% decline in wages on an annualized basis (or about 25.6% and 3.2% overall, respectively) over the 1980-2012 period.

Next, we examine outcomes at the industry level. To do so, we use industry employment shares to collapse our innovation measure at the industry level. Breakthrough patents that are most similar to the industry's workforce are strongly positively predictive of industry productivity and output, yet they predict declines in the labor share of output. The magnitudes are sizeable: a one-standard deviation increase in our technology indicator is associated with a 1.5 percent increase in measured productivity per year, but a 0.6 percent decline in the labor share of production workers. Given that the improved production methods our measure captures are part of 'intangible' capital, our evidence supports the view that much of the recent decline in the labor share has been due to the rising importance of

the last strongholds of specialized handicraft skill, has undergone a technological revolution resulting in the almost complete disappearance of the hand branch of the industry and the elimination of two skilled trades and one semiskilled, and also the partial elimination of the skilled flatteners."

intangible capital.

In brief, our measure of occupational exposure is largely associated with higher measured productivity but simultaneously with employment and wage declines. Thus, it appears that our procedure mostly identifies the exposure of occupations to labor-saving innovations. We should emphasize that this fact does not mechanically follow from our methodology, which compares the similarity of the occupation's task descriptions to a given innovation. Certainly, many of the innovations in these period likely complement the skills of existing workers; hence our findings could understate the actual degree to which technological innovation has displaced workers. To quantify this possibility, in the last part of the paper we construct a statistical textual factor designed to maximize the in-sample predictability of employment declines. Our baseline innovation measures have a correlation of approximately 75% with this statistical factor; excluding patents in computers and software increases this correlation to 92%. We conclude that our methodology mostly identifies labor-displacing innovations—especially in the pre-1980 period.

Our work complements several strands of the literature. Existing work has emphasized the complementarity between technology and certain types of worker skills (Goldin and Katz, 1998, 2008; Autor, Levy, and Murnane, 2003; Autor, Katz, and Kearney, 2006; Goos and Manning, 2007; Autor and Dorn, 2013); or the substitution between workers and new forms of capital (Hornstein, Krusell, and Violante, 2005, 2007; Acemoglu and Restrepo, 2020). Most of these papers do not distinguish between innovation and adoption. By contrast, our use of patent data implies that we necessarily focus on innovation rather than adoption of existing technologies.

Perhaps closest to our work is the large literature documenting the secular decline in occupations specializing in routine tasks, starting in the late 20th century (Autor and Dorn, 2013; Acemoglu and Autor, 2011; Autor et al., 2003). The key idea is that routine tasks can be easily codified into a sequence of instructions. Hence, such tasks are relatively more prone to labor-saving technological change than other more complex tasks. Put differently, an occupation's routine task intensity (RTI) measures its exposure to automation. Despite the success that this literature has had in explaining which occupations have been exposed to technologies, and what have been the effects, it is still an open question how this exposure changes over time, which technologies relate to which types of tasks and which occupations, and whether or not technological unemployment is a robust phenomenon in other time periods. We contribute to these questions by providing a text-based measure of time-varying occupation similarity to new innovations that spans over 150 years (from 1850-2010), allowing us to examine the changing task content of innovations and the subsequent occupational outcomes over a longer time period than has previously been studied. In this respect, our work is close to Atack, Margo, and Rhode (2019) who analyze how workers' task transitioned from hand to machine production in the late 19th century.

Last, our work is also closely related to recent work by Webb (2019), who also analyzes the similarity between patents and occupation task descriptions. Our work differs in methodology and scope. In terms of methodology, Webb (2019) measures this similarity using the co-occurrence of verb-object pairs in the patent title and abstract of patents with verb-object pairs in the job task descriptions. He uses word hierarchies obtained from WordNet to determine similarity in verb-object pairings. By contrast, we infer document similarity by using geometric representations of word meanings (Global vectors for word representation, henceforth GloVe) that have been estimated directly from word co-occurrence counts; we use not only the abstract but the entirety of the patent document—which includes the abstract, claims, and the detailed description of the patented invention. More importantly however, our work differs in scope and aim. Webb (2019) focuses on automation and the future of work, and thus restricts attention to patents identified as being related to robots, AI, or software. As a result, the analysis in Webb (2019) is largely cross-sectional in nature as he focuses on a single technological episode—the rise of AI and robots. In contrast to Webb (2019), we have a much broader focus: we are interested in constructing time-series indicators to understand the relation between innovation and employment over different technological episodes. As such, we compute occupation-patent distance measures for all occupations and the entire set of USPTO patents dating to the beginning of the Second Industrial Revolution. As a result, our analysis primarily captures manufacturing patents, as opposed to robotics, software, or AI.

The rest of the paper is organized as follows. Section 1 describes our methodology for inferring the similarity between patent documents and occupations. Section 2 constructs time-series indices of occupations' exposure to innovation. Section 3 relates our time-series indices to occupational outcomes—employment and wages. Last, Section 4 concludes.

1 Similarity of Patents and Job Descriptions

Here, we describe our methodology in identifying which patents are related to which occupations. Section 1.1 describes our data sources; Section 1.2 describes our methodology for estimating the similarity of patent documents to job descriptions from the Dictionary of Occupational Titles (DOT).

1.1 Data Construction

We obtain text data for measuring patent/job task similarity from two sources. Job task descriptions come from the revised 4th edition of the Dictionary of Occupation Titles (DOT) database. We use the patent text data parsed from the USPTO patent search website in Kelly et al. (2020), which includes all US patents beginning in 1976, comprising patent numbers 3,930,271 through 9,113,586, as well as patent text data obtained from Google patents for pre-1976 patents. Our analysis of the patent text combines the claims, abstract, and description section into one patent-level corpus for each patent. Since the DOT has a very wide range of occupations (with over 13,000 specific occupation descriptions) we first crosswalk the DOT occupations to the considerably coarser and yet still detailed set of 6-digit occupations in the 2010 edition of O*NET. We then combine all tasks for a given occupation at the 2010 O*NET 6-digit level into one occupation-level corpus. See Appendix 5 for further details on cleaning and preparing the text files for numerical representation.

1.2 Measuring Textual Similarity

Our goal is to construct a distance measure between a patent document and a job description. The main challenge is to identify meaningful connections between two sets of documents. A common approach for computing document similarity would be to create a matrix representation of each document, with columns representing document counts for each term (or some weighting of term counts) in the dictionary of all terms contained in the set of documents, and with rows representing each document. Similarity scores could then be computed simply as the cosine similarity between each vector of weighted or unweighted term counts:

$$\rho_{i,j} = \frac{V_i}{||V_i||} \cdot \frac{V_j}{||V_j||} \tag{1}$$

Here V_i and V_j denote the vector of potentially weighted terms counts for documents i and j. This is commonly called the bag-of-words approach, and has been used successfully in many settings.

A variant of this approach is used by Kelly et al. (2020), who construct measures of patent novelty and impact using as inputs pairwise similarity comparisons between patents. Since patent documents have a structure and a legalistic vocabulary that is reasonably uniform, this approach works quite well for patent-by-patent comparisons. However, in our case the two documents come from different sources: job task text has a very different structure and vocabulary from the patent text. Since V_i and V_j are highly sparse vectors with most elements equal to zero, the patent and task term vectors are likely to be nearly orthogonal. What's more, (10) has no way of accounting for words with similar meanings. Consider a set of two documents, with the first document containing the words "dog" and "cat" and the other containing the words "puppy" and "kitten". Then the count representation of the first document is $V_1 = [1, 1, 0, 0]$ and $V_2 = [0, 0, 1, 1]$ so $\rho_{1,2} = 0$ even though they carry essentially the exact same meaning.

We address this issue of synonyms by leveraging recent advancements in natural language processing that allow for accurate dense vector representations of word meanings.[4] In particular, we use word vectors provided by Pennington et al. (2014) that were trained on 42 billion word tokens of web data from Common Crawl.[5] This contains a vocabulary of 1.9 million 300-dimensional vector representations of word meanings, commonly called word embeddings. In order to move from word vectors to document vectors, we simply take a weighted average of the word embeddings within a document. Denote by A_i the set of word vectors for document i and X_i the document vector of a patent or occupation text i

$$X_i = \sum_{x_k \in A_i} w_{i,k} x_k \tag{2}$$

Here $w_{i,k}$ is the term-frequency-inverse-document-frequency (TFIDF) defined as

$$w_{i,k} \equiv TF_{i,k} \times IDF_k \tag{3}$$

where

$$TF_{i,k} = \frac{c_{i,k}}{\sum_j c_{i,j}} \tag{4}$$

with $c_{i,j}$ denoting the count of the jth word in document i. Thus the term-frequency is the count of word k in document i divided by the number of terms in document i. The inverse-document frequency is

$$IDF_k = \log\left(\frac{\text{\# of documents in sample}}{\text{\# of documents that include term } k}\right) \tag{5}$$

[4]The two most popular are the "word2vec" method of Mikolov, Sutskever, Chen, Corrado, and Dean (2013) and the global vectors for word representation introduced by Pennington, Socher, and Manning (2014). These papers construct mappings from extremely sparse and high-dimensional word co-occurence counts to dense and comparatively low-dimensional vector representations of word meanings called word embeddings. Their word vectors are highly successful at capturing synonyms and word analogies (vec(king) − vec(queen) ≈ vec(man) − vec(woman) or vec(Lisbon) − vec(Portugal) ≈ vec(Madrid) − vec(Spain), for example). Thus they are well-suited for numerical representations of the "distance" between words.

[5]These are available for download at https://nlp.stanford.edu/projects/glove/.

Thus $TFIDF_{i,k}$ overweighs word vectors for terms that occur relatively frequently within a given document and underweighs terms that occur commonly across all documents. The use of TF-IDF weighting is common in natural language processing. We compute the inverse-document-frequency for the set of patents and occupation tasks separately, so that patent document vectors underweight word embeddings for terms appearing in many patents and occupation vectors underweight word embeddings for job task terms that appear in the task descriptions of many other occupations.

To understand how our metric differs from the standard bag-of words approach it is useful to briefly examine how word embeddings are computed in Pennington et al. (2014). Denote the matrix X as a $V \times V$ matrix of word co-occurence counts obtained over a set of training documents, where V is the number of words in the vocabulary. Then $X_{i,j}$ tabulates the number of times word j appears in the context of the word i.[6] Denote $X_i = \sum_k X_{i,k}$ as the number of times any word appears in the context of word i, and the probability of word j occuring in the context of word i is $P_{i,j} \equiv X_{i,j}/X_i$. The goal of the word embedding approach is to construct a mapping $F(\cdot)$ from some d-dimensional vectors x_i, x_j, and $\tilde{x}_k$ such that

$$F(x_i, x_j, \tilde{x}_k) = \frac{P_{i,k}}{P_{j,k}} \tag{6}$$

Imposing some conditions on the mapping $F(\cdot)$, they show that a natural choice for modeling $P_{i,k}$ in (6) is

$$x_i^T \tilde{x}_k = \log(X_{i,k}) - \log(X_i) \tag{7}$$

Since the mapping should be symmetric for i and k they add "bias terms" (essentially i and k fixed effects) which gives

$$x_i^T \tilde{x}_k + b_i + b_k = \log(X_{i,k}) \tag{8}$$

Summing over squared errors for all pairwise combinations of terms yields the weighted least squares objective

$$\text{Min}_{x_i,\tilde{x}_k,b_i,b_k} \sum_{i=1}^{V} \sum_{j=1}^{V} f(X_{i,j}) \left(x_i^T \tilde{x}_k + b_i + b_k - \log(X_{i,j})\right)^2 \tag{9}$$

Here the observation-specific weighting function $f(X_{i,j})$ equals zero for $X_{i,j} = 0$ so that the log is well defined, and is constructed to avoid overweighting rare occurences or extremely

[6] Pennington et al. (2014) use a symmetric 10 word window to determine "context" and weight down occurences that occur further away from the word (one word away receives weight 1, two words away receives weight 1/2, etc.).

frequent occurences. The objective (9) is a highly-overidentified least squares minimization problem. Since the solution is not unique, the model is trained by randomly instantiating x_i and $\tilde{x}_k$ and performing gradient descent for a pre-specified number of iterations, yielding d-dimensional vector representations of a given word. Here d is a hyper-parameter; Pennington et al. (2014) find that $d = 300$ works well on word analogy tasks.

Since (9) is symmetric it yields two vectors for word i, x_i and $\tilde{x}_i$, so the final word vector is taken as the average of the two. The ultimate output is a dense 300-dimensional vector for each word i that has been estimated from co-occurence probabilities and occupies a position in a word vector space such that the pairwise distances between words (i.e. using a metric like the cosine similarity) are related to the probability that the words occur within the context of one another and within the context of other similar words. Note that the basis for this word vector space is arbitrary and has no meaning; distances between word embeddings are only well-defined in relation to one another and a different training instance of the same data would yield different word vectors but very similar pairwise distances between word vectors.

Our method for backing out a geometric representation of the "meaning" of a document in (2) is to construct a weighted average of the meaning of all words in the document. Thus our vector representation of documents retains the 300-dimensional structure of the individual word constituents; these vectors are much denser and smaller than the very large and sparse document vectors in the standard bag of words methodology. In brief, there are two key characteristics that differentiate our approach relative to bag of words techniques. First, X_i is no longer a sparse vector like V_i. Moreover, because of the way word vectors are estimated, our method allows vectors containing similar words to be "close" to one another. Thus, relative to the bag of words approach our method: (1) constitutes a large dimensionality reduction; and, (2) can incorporate a notion of synonyms/distances between word meanings.

Armed with a vector representation of the document that accounts for synonyms, we next use the cosine similarity to measure the similarity between patent i and occupation j:

$$\text{Sim}_{i,j} = \frac{X_i}{||X_i||} \cdot \frac{X_j}{||X_j||} \tag{10}$$

This is the same distance metric as the bag of words approach, except now X_i and X_j are dense vectors carrying a geometric interpretation akin to a weighted average of the semantic meaning of all nouns and verbs in the respective documents.

Returning to our previous example of two documents with the words "dog; cat" and "puppy; kitten" , these documents are now represented as

$$X_1 = (1/2) \times \log(2) x_{dog} + (1/2) \times \log(2) x_{cat} \tag{11}$$

and similarly for X_2.[7] Here x_{dog}, x_{cat} would have been trained using the Pennington et al. (2014) method described above on a very large outside set of documents. Hence, in this case since word vectors are estimated such that $x_{dog} \approx x_{puppy}$ and $x_{cat} \approx x_{kitten}$, we now have $\text{Sim}_{1,2} \approx 0.81$ using the word vectors estimated by Pennington et al. (2014). A weighted average word embedding approach has been shown in the natural language processing literature to achieve good performance on standard benchmark tests for evaluating document similarity metrics relative to alternative methods that are much more costly to compute (see, e.g. Arora, Liang, and Ma, 2017). A relative disadvantage is that it ignores word ordering—which also applies to the more standard 'bag of words' approach for representing documents as vectors. However, since we have dropped all stop words and words that are not either a noun or a verb, retaining word ordering in our setting is far less relevant.

Our methodology bears brief comparison with recent work by Webb (2019), who also analyzes the similarity between a patent and O*NET job tasks. Webb (2019) focuses on similarity in verb-object pairs in the title and the abstract of patents with verb-object pairs in the job task descriptions and restricts his attention to patents identified as being related to robots, AI, or software. He uses word hierarchies obtained from WordNet to determine similarity in verb-object pairings. By contrast, we infer document similarity by using geometric representations of word meanings (GloVe) that have been estimated directly from word co-occurence counts. In addition to employing a different methodology, we also have a broader focus: we compute occupation-patent distance measures for all occupations and the entire set of USPTO patents since 1836. Furthermore, we use not only the abstract but the entirety of the patent document—which includes the abstract, claims, and the detailed description of the patented invention.

1.3 Examples

A potential downside on relying on textual similarity measures is that their effectiveness can be highly dependent on a specific context. To illustrate the effectiveness of our measure, we consider a few representative examples.

A key advantage of our measure is that it is available over long periods of time, and thus

[7] This comes by noticing that in our simple example, $TF_{1,dog} = 1/2$ and $IDF_{dog} = \log(2)$, with similar logic applying to "cat"; this proceeds analogously for document 2 containing "puppy" and "kitten".

allows us to study very different technologies from three distinct periods of technological change–the Second Industrial Revolution of the late 1800's, the period spanning the from 1920s to around 1940, and the information technology revolution spanning the end of the 20th and beginning of the 21st centuries. For example, consider patent 276,146, titled "Knitting Machine", issued in the height of the Second Industrial Revolution in 1883. The occupation that is most closely related to this patent is "Textile Knitting and Weaving Machine Setters, Operators, and Tenders"; the next most similar occupation is "Sewing Machine Hand Operators", followed by "Sewers, hand". Next consider the patent for "Metal wheel for vehicles (1,405,358), which is issued in 1922. The occupation most closely related to this patent is "Automotive Service Technicians and Mechanics", with other production and metal machine workers following. Finally, we examine a patent from a very different era and representing a very different technology. The patent, entitled "System for managing financial accounts by a priority allocation of funds among accounts," is U.S. patent number 5,911,135 and was issued in 1999. The top occupations? Financial managers, credit analysts, loan interviewers and clerks, and so on. The relevance of this patent to the most similar occupations is readily apparent. All three patents are in the list of breakthrough patents identified by Kelly et al. (2020).

We next perform the reverse exercise, where we fix a particular occupation, and list the most relevant patents. The occupations we choose are cashiers, loan interviewers and clerks, and railroad conductors. Table 2 lists the top five patents that are linked to each of these occupations. Examining the patent tiles, we see that each one of these patents is directly related to the work performed by the given occupation. For example, one of the top patents for cashiers is "Vending type machine dispensing a redeemable credit voucher upon payment interrupt" (patent 5,055,657); the top patent for loan interviewers and clerks is titled "Automatic business and financial transaction processing system" (patent number 6,289,319). And finally, for rail road conductors, titled "Automatic train control system and method" (patent 5,828,979) is the top patent. In general the patents showing up on this list represent technologies that (1) relate to the work performed by individuals in that the occupation; and (2) if adopted, appear likely to be able to change the way that an occupation performs its core work functions and/or substitute for work done by that occupation.

In sum, these examples illustrate the ability of our method in identifying technologies that are related to a particular occupation. However, it is not immediately obvious whether these technologies benefitted workers in these occupations or whether they led to the displacement of workers. Most likely, some of these technologies benefitted some workers at the expense of others. To illustrate the potential for such reallocation, we consider two examples of labor saving technologies from Jerome (1934). First, consider two key innovations in the textile weaving industry during the early 20th century, the Barber-Colman warp-tying machine (patent 1,115,399) and the drawing-in machine (patent 1,364,091). Both of these technologies benefitted skilled workers at the expense of unskilled labor. Jerome (1934) notes that, the Barber-Colman warp-tying machine "will do the work of about 15 hand operators" while "it can be run by one tender." Similarly, he notes that "It is estimated that each (drawing-in machine) machine, requiring ordinarily the attention of one operator and half the time of an assistant, replaces from 5 to 6 hand drawers-in." Both of these patents are identified as breakthrough patents by Kelly et al. (2020). In terms of related occupations, our methodology identifies various types of textile workers as being the some of the most relevant.

However, not all labor-saving technologies benefit skilled labor. For instance, consider two major innovations in the window glass industry during the late 19th century—the Colburn sheet machine (patent 840,833) and the cylinder machine (patent 814,612). Following their introduction, the manufacturing process for window glass switched from being hand-made to being entirely mechanized by 1925. The displacement of skilled workers was rapid: by 1905 many hand plants had gone out of business, wages of blowers and gatherers were reduced 40 per cent. Jerome (1934) summarizes their impact thus: "In the quarter century following the introduction of machine blowing, the window-glass industry, one of the last strongholds of specialized handicraft skill, has undergone a technological revolution resulting in the almost complete disappearance of the hand branch of the industry and the elimination of two skilled trades and one semiskilled, and also the partial elimination of the skilled flatteners. The contest for supremacy now lies between the cylinder and the sheet machine processes." Both of these patents are in the top 10% of the Kelly et al. (2020) measure. In terms of our methodology, we identify "glaziers" and "molders, shapers, and casters, except metal and plastic" as being among the most related to these two patents.[8]

These two examples illustrate that the impact of a new technology on a given worker is not ex-ante obvious. Some technologies may replace un-skilled workers, while others may displace highly specialized and skilled workers. Indeed as Jerome (1934) notes, glass workers displaced by the sheet and cylinder machines in their time were considered to be members of skilled trade. Further, new technologies may also generate demand for new skills—for example, the operators of the Barber-Colman warp-tying machine—hence their long-run effects may be different than their short-run impact.

[8]The occupation titled "molders, shapers, and casters, except metal and plastic", which corresponds SOC code 519195, has a sub-occupation called "glass blowers, molders, benders, and finishers".

2 Measuring Technology Exposure over Time

Our analysis so far delivers a measure of similarity between a given patent and a given occupation. Here, we focus our attention on constructing time-series indices of technological exposure at the occupation level.

2.1 Methodology

The first challenge in constructing a time-varying index lies in choosing how to appropriately identify the 'amount' of innovation that occurs at a given point in time. One possibility would be to count the number of patents; however, this approach is unlikely to be fruitful, since not all patents are equally important (see, e.g. Hall, Jaffe, and Trajtenberg, 2005). Various approaches have been proposed, which essentially weight patents by their forward citations (Hall et al., 2005); estimates of their market value (Kogan, Papanikolaou, Seru, and Stoffman, 2017); or their textual similarity to prior and subsequent patents (Kelly et al., 2020).

We choose the Kelly et al. (2020) approach for two reasons: first, unlike forward citations, their measure is available for the entirety of our sample; second, we are primarily interested in the contribution of a patent on the technology frontier rather than their private value to their firm. Kelly et al. (2020) identify breakthrough innovations as those that are both novel (whose descriptions are distinct from their predecessors) and impactful (they are similar to subsequent innovations). Figure 2 plots their time-series index of breakthrough innovations—those that lie above the 90th percentile in the ratio of forward to backward textual similarity.

We combine the occupation by patent text similarities with the breakthrough patent index to create our time-varying measure of occupational exposure to new technologies. We define our time series index of exposure of occupation i to technology at time t as

$$\eta_{i,t} = \frac{1}{\kappa_t} \sum_{j \in \Gamma_t} \tilde{\rho}_{i,j} \times \mathbf{1}(\tilde{q}_{j,t} \geq \tilde{q}_{p90}). \tag{12}$$

Our time-series index (12) aggregates our patent-occupation similarity scores across all breakthrough patents issued in year t. Specifically, we sum over occupation exposures $\tilde{\rho}_{i,j}$ across patents $j \in \Gamma_t$ that are issued in year t. We restrict attention to breakthrough patents, that is, patents whose Kelly et al. (2020) ratio of importance $\tilde{q}_{j,t}$ exceeds the 90th percentile $\tilde{q}_{p90}$. When computing $\tilde{\rho}_{i,j}$ in (12), we perform the following adjustments to our raw occupation-patent exposure $\rho_{i,j}$. First, we remove the interaction of issue year and technology fixed effects.[9] Second, we impose sparsity: after removing the fixed effects we

[9]We do so in order to account for language and structural differences in patent documents over time and

set all patent $\times$ occupation pairs to zero that are below the 80th percentile in fixed-effect adjusted similarity. Last, we scale the remaining non-zero pairs such that a patent/occupation pair at the 80th percentile of tech-class $\times$ year adjusted similarities has a score equal to zero and the maximum adjusted score equals one.

Table 3 lists the top and bottom five occupations by average exposure over the time period spanning 1850 to 2002. The most exposed occupation is titled "Inspectors, Testers, Sorters, Samplers, and Weighers". The top occupations tend to be those working in production and manufacturing type jobs, which are commonly posited to be among the type of occupations most affected by new technologies. The least exposed occupations are mental health counselors, dancers, funeral attendants, judges, and clergy, all representing service job types that are unlikely to have the nature of their work substantially changed by new technologies.

2.2 Exposures of tasks to technological change

One way to summarize our time-series indices is to examine what kinds of tasks are performed by occupations that are mostly related to technological change at different points in time. We use the six task categories constructed by Acemoglu and Autor (2011), which include non-routine cognitive (analytical), non-routine cognitive (interpersonal), non-routine manual (interpersonal) non-routine manual (physical), routine cognitive, and routine manual tasks. In the exercises that follow we fix the distribution of tasks using SOC labor supply weights provided by Acemoglu and Autor (2011). They scale the task measures so that each task type has a cross-sectional distribution of mean 0 and standard deviation 1.

Table 4 shows the correlations of our measure $\eta_{i,t}$ with each task type. Occupations most exposed to innovation tend to score higher on non-routine manual (physical) and routine manual tasks and score low on non-routine cognitive (interpersonal) and non-routine manual (interpersonal) tasks. This is consistent with the pattern observed in Table 3, with service-type occupations that specialize in person-to-person interaction scoring especially low on average exposure.

We next examine how this association has changed over time. Let $T_{i,w}$ index the task score for task type w and occupation i. Our time-varying measure of task similarity is then

technology areas. Patents have become much longer and use much more technical language over the sample period, and the OCR text recognition of very early patents is far from perfect. Specific technology classes also use very different language from one another (for example, patents for new chemical compounds have a disparate vocabulary from patents on new manufacturing tools). Removing year $\times$ technology class fixed effects helps to adjusts for these differences. For consistency, we also similarly modify the Kelly et al. (2020) procedure: rather than removing year fixed effects to identify the patents in the right tail of the unconditional distribution, we remove the interaction between year and technology fixed effects.

given by

$$\lambda_{w,t} = \sum_i \eta_{i,t} \times T_{i,w} \times \omega_i \tag{13}$$

Here ω_i are the Acemoglu and Autor (2011) labor supply shares. Because the ω_i sum to one and $T_{i,w}$ has mean zero and standard deviation one when weighted by ω_i, the measure $\lambda_{w,t}$ is the coefficient from a cross-sectional regression of $\eta_{i,t}$ on $T_{i,w}$ weighted by ω_i. Thus positive values of $\lambda_{w,t}$ indicate that $\eta_{i,t}$ is correlated with occupations that have high scores on a given task type, and negative values indicate that $\eta_{i,t}$ is associated with occupations having below average scores for that particular task.

In Figure 3 we plot the time series of $\lambda_{w,t}$ for all six task types. Consistent with Table 4 we see interpersonal tasks consistently ranking at the bottom and non-routine manual (physical) and routine manual tasks at the top. Of note are the three breakthrough innovation waves documented previously. Whereas the innovation waves of the 1880s and 1920s-1930s merely exacerbated this pattern, with upward shifts in non-routine manual (physical)/routine manual tasks running parallel with downward shifts in interpersonal tasks and non-routine cognitive (analytical) tasks. However, the post-1980 period of innovation is distinctly different. In this period routine cognitive and non-routine cognitive (analytical tasks) have increased in their association with innovations. Meanwhile, the decreased association of recenet innovations with interpersonal is even more pronounced in this time period.

This pattern is driven by the changes in the composition of innovations within the information/communication and electronics technology classes, as well as the increased importance of these technologies for aggregate innovation. We show this by constructing four technology-class specific measures of $\eta_{i,t}$ for information and communications technologies, electronics, manufacturing process, and construction/engineeering patents. These technology classes are constructed so as to have no overlap.[10]

Figures 4 and 5 plot the time series for information/communication and electronics patents. Note the stark rise in the association with routine-cognitive and non-routine cognitive (analytical) tasks among these two technology class types. This pattern is driven by information technology revolution that has led to the modern digitalization of the workplace. Occupations that relate to these type of innovations have a distinctly different task profile than the most prevalent technologies of past innovation waves.

In Figures 6 and 7 we show the same plots for manufacturing process versus construction and engineering patents, respectively. No such distinctive pattern can be found among these

[10] We follow Kelly et al. (2020) in using the CPC patent technology classification scheme to assign patents into 12 broad technology classes.

technology classes, and in fact the recent innovation wave has behaved in essentially the exact same manner as those in the past. Thus the composition of task content within manufacturing and construction type innovations has not changed in response to the recent IT revolution; other technology classes have instead become more prevalent and strongly associated with cognitive tasks. Additionally, at no point in time or within any technology class have the non-routine manual (interpersonal) tasks ceased their strong negative association with the new technologies described in patent text. This makes sense in the context of the findings of Autor and Dorn (2013), who show that service occupations have increased in importance at the expense of occupations heavily exposed to automation, and also Deming (2017), who documents an increased importance of social skills in the labor market.

We next examine the types of occupation that are exposed to innovation over time. We group each occupation into one of eight broad occupation group types: service; sales and office; production, transportation, and material moving; natural resources, construction, and maintenance; management, business, and financial; healthcare practitioners; education, legal, community service, arts and media; and, computer, engineering, and science. Within each of these groups we take the average of $\eta_{i,t}$ and then scale across the eight groups each year so that the total sums to one. The results are plotted in Figure 8, which, shows that while production and construction type occupations have consistently been the most exposed, there has been a consistent upward rise in the importance of computer, engineering, and science occupations over time. Interestingly, this rise became more pronounced in the decades immediately preceding the IT revolution, starting around 1950 and increasing in prevalence ever since. Sales and office occupations have also seen an increased relationship with innovation, as well as management/business occupations, though these two groups remain small in their overall exposure. Also of note is that service occupations have never had a large relationship with innovation and the fraction has hardly changed over time.

3 Technology exposures and labor market outcomes

Armed with a measure of technology exposure across occupations and time, we next examine the relation between innovation and labor market outcomes. To motivate our analysis, we first present a simple model. We then examine the correlation between our occupation-specific measures of technology exposure and labor market outcomes.

3.1 A model of job displacement

Here, we present a simple model of how technological progress affects labor income. The basic unit of production in the economy is a task $i \in (0,1)$; aggregate output is a composite of all tasks,

$$Y = \int_0^1 y(i)^{\frac{\sigma-1}{\sigma}}\, di, \qquad \text{where} \quad \sigma > 1. \tag{14}$$

Given (14), optimization by the final goods firms implies a demand curve for each task,

$$\frac{\sigma-1}{\sigma}\, y(i)^{-\frac{1}{\sigma}} = p(i). \tag{15}$$

Each task can be performed either by labor l or by machines or automated processes k. Labor and automation are perfect substitutes at performing each task,

$$y(i) = \gamma(i)\, l(i) + \xi(i)\, k(i). \tag{16}$$

Here, $\gamma(i)$ is the productivity of labor in task i and $\xi(i)$ is the productivity of machines/automation in performing task (i).

Perfect competition in the production of tasks implies that marginal revenue is equated to marginal costs,

$$p(i) = \min\left(\frac{w(i)}{\gamma(i)}, \frac{q(i)}{\xi(i)}\right) \tag{17}$$

where $w(i)$ and $q(i)$ is the compensation of labor and machines, respectively. Given that labor and automation are perfect substitutes, some tasks will be performed only by labor, while others only by machines. As a result,

$$p(i)\,\gamma(i) = W(i) \quad \text{if} \quad \frac{W(i)}{\gamma(i)} < \frac{q(i)}{\xi(i)} p(i)\,\xi(i) = q(i) \quad \text{if} \quad \frac{W(i)}{\gamma(i)} > \frac{q(i)}{\xi(i)} \tag{18}$$

The supply of automation services is perfectly elastic—that is, the economy can convert one unit of output Y to one unit of automation services $k(i)$ at a constant price of $q(i) = 1$. In contrast to automation, however, the supply of skilled labor is inelastic (at least in the short run). In particular, performing a task requires a specific skill, and skills are fully task-specific. We denote by $L(i)$ the total supply of skilled labor for task i. Since skills are task-specific, and hence cannot be reallocated across tasks, labor will participate in the production of automated tasks, albeit at a depressed wage. That is, labor compensation for each task will satisfy

$$w(i) \leq \frac{\gamma(i)}{\xi(i)}. \tag{19}$$

We next solve for labor compensation in equilibrium. Consider the tasks for which labor is either sufficiently productive relative to automation such that

$$w(i) < \frac{\gamma(i)}{\xi(i)}. \tag{20}$$

For the task to be performed only by labor, we need that skilled labor is either sufficiently productive relative to machines or the supply of skills is high,

$$\frac{\sigma}{\sigma - 1}\Big(\gamma(i)\, L(i)\Big)^{\frac{1}{\sigma}} > \xi(i). \tag{21}$$

For these tasks, labor compensation is only a function of the productivity of labor and the supply of skills,

$$w(i) = \frac{\sigma - 1}{\sigma}\gamma(i)\Big(\gamma(i)\, L(i)\Big)^{-\frac{1}{\sigma}} \tag{22}$$

By contrast, tasks for which the inequality in (21) does not hold will be partially automated. For these tasks, the productivity of automation directly affects wages, since $w(i) = \gamma(i)/\xi(i)$.

Absent any other modification, the model always generates full employment, since workers are willing to work for any wage. Thus, we introduce a minimum level of compensation, $\underline{w}$, below which workers will not work in production. The most straightforward interpretation is that the skills relevant for task i can also be used in home production, which yields a benefit of $\bar{w}$. As a result, some tasks will be fully automated—these are tasks for which the equilibrium level of compensation falls below the threshold $\underline{w}$, or equivalently,

$$\underline{w} \geq \min\left[\frac{\gamma(i)}{\xi(i)}, \frac{\sigma - 1}{\sigma}\gamma(i)\Big(\gamma(i)\, L(i)\Big)^{-\frac{1}{\sigma}}\right] \tag{23}$$

Figure 1 summarizes the equilibrium level of automation, labor compensation and employment as a function of γ and ξ. We see that increases in the productivity of automation in a given task $\xi(i)$ leads to increases in the likelihood that the task is either partially or fully automated. In the region where the task is partially automated, increasing $\xi(i)$ also depresses wages. If $\xi(i)$ is sufficiently high then the task is fully automated, as the equilibrium level of compensation falls below the workers' outside option $\underline{w}$. Here, we note that, unlike the Acemoglu-Restrepo model, improvements in automation displaces worker skills both at the intensive as well as the extensive margin. That is, an increase in $\xi(i)$ both increases the likelihood that a given task will be (partially) automated, but also depresses labor compensation for already (partially) automated tasks. This difference is driven by our assumption that labor skills are specific to a given task and cannot be reallocated. Since the supply of labor skills is inelastic, wages decrease as the productivity of automation increases.

We next turn to the compensation for a given worker/occupation. An occupation j is a finite collection of tasks $I(j)$. A worker n has a given set of task-specific skills $s(i,n)$ that is relevant to her occupation. Here we equate workers with an occupation: workers are born to a given occupation, so there is no possibility to retrain or acquire new skills. Given these assumptions, the compensation of worker n is equal to

$$W(n) = \sum_{i \in I(j)} s(i,n)\, w(i). \tag{24}$$

The total supply of skills is equal to

$$L(i) = \int_n s(i,n) dn \tag{25}$$

The model thus links technological innovation, which can manifest either as improvements in labor productivity or the productivity of automation, and worker wages. To illustrate the mechanics of the model, consider an occupation that requires two tasks i_0 and i_1. The first task is one that not automated ($\xi(i)$ is quite low), while the second is partially automated already. Increases in the productivity of automation in task i_1 lower equilibrium wages,

$$\frac{d \log W(n)}{d \log \xi(i_1)} = -\frac{s(i_1,n)\, w(i_1)}{s(i_0,n) w(i_0) + s(i_1,n)\, w(i_1)} \tag{26}$$

Conversely, increases in labor productivity γ increase wages

$$\frac{d \log W(n)}{d \log \gamma(i_1)} = \frac{s(i_1,n)\, w(i_1)}{s(i_0,n) w(i_0) + s(i_1,n)\, w(i_1)} \tag{27}$$

Here, we note these are local effects that are relevant in the region of partial automation. Larger increases in $\xi(i)$ or $\gamma(i)$ can push task i outside the region of partial automation. In particular, a larger increase in $\xi(i_1)$ can lead to the task being fully automated, and therefore to a discontinuous decline in wages. In addition, the magnitude of these effects on worker wages depends on workers' skill weights. All else equal, more skilled workers will experience a larger decline in wages in response to automation.

3.2 Effects of Innovation Exposures On Employment and Wages

Having documented which occupations and tasks are linked to innovations over the last 150+ years, we now proceed to examine the subsequent outcomes of exposed occupations.

The model in section 3.1 outlines how improvements in capital that directly substitute for worker skills can reduce the wages and employment for workers whose occupational tasks are exposed. In that model the supply elasticities of labor and automation technologies were fixed so as to isolate the displacement effect that new technologies have on occupations for whose tasks they substitute. However, in a more general model it is not obvious that labor demand should unambiguously decrease following an increase in the productivity of capital that directly substitutes for labor. Acemoglu and Restrepo (2018) show that new capital also induces a productivity effect that leads to increased labor demand because it raises the demand for all inputs. On the other hand, the displacement effect due to the increased productivity of capital relative to the labor for which it substitutes has a clear negative effect on labor demand.

It is ultimately an empirical question whether or not the job displacement highlighted by our model dominates potential productivity effects. Note also that although our measure is intended to pick up innovations that substitute for the work performed by specific occupations, we may also pick up technologies that are more likely to represent Hicks-neutral productivity shocks which improve the marginal productivity for labor and hence do not reduce labor demand. Thus it is not clear whether finding a positive marginal effect of innovation on labor comes because substituting innovations raise overall productivity and this effect dominates or whether innovation primarily complements for rather than substitutes for labor. Negative effects on labor demand, on the other hand, must be driven by the displacement channel outlined in our simple model.

We perform several empirical exercises to examine the relationship between our measure and subsequent labor market effects. First we look at the average relationship between innovation exposure and subsequent growth in the employment shares of occupations.

3.2.1 Innovation exposure and employment growth in the pre-1980 period

We gather Census data from IPUMS and compute aggregate employment shares for occupations in Census years spanning 1850-2010. We use the 1950 Census occupation definition for pre-1950 Census years since the more updated 1990 Census classification scheme is only available in post-1950 Census years. We make use of the 1990 Census occupation classifications for the years they are available. We then crosswalk Census occupations to the David Dorn occ1990dd classification scheme using the crosswalk files provided on his website and aggregate our measure $\eta_{i,t}$ to the occ1990dd-level by averaging across 6-digit SOC codes within an occ1990dd code. This results in a Census-year by occ1990dd panel of occupation employment shares. We then run the panel regression

$$\log(Y_{i,t+k}) - \log(Y_{i,t}) = \alpha_0 + \alpha_t + \beta\eta_{i,t} + \lambda Y_{i,t} + \epsilon_{i,t} \tag{28}$$

for $k = 10, 20$ years and for Census years spanning from 1850-2010. Here $Y_{i,t}$ is the employment share in total non-farm employment.[11] In Table 5 we report the results. The left panel reports the estimates for the full sample and for both the 10- and 20-year horizons, while the right panel reports results for 1850-1920 and 1930-1990 subsamples at the 20-year horizon.[12] Observations are weighted by the employment share of the given occupation and standard errors are clustered by occupation. The employment growth rates are expressed in annualized percentage terms and the innovation exposures are standardized. We also include time fixed effects to average out the changes in aggregate innovation over time and focus on employment responses to the cross-sectional dispersion in innovation exposure.

Consistent with the displacement channel, we see that the point estimates are strongly negative for the whole sample, at both time horizons, and for both of the subsamples considered. Since the growth rates are in annualized percentages and $\eta_{i,t}$ is standardized, the coefficients imply that a standard deviation increase in $\eta_{i,t}$ is associated with a -.41% annualized decline in employment over the next 10 years and a -0.70% percent decrease in employment over the next 20 years. These effects are strongly significant, with t-stats of -4.18 and -4.59, respectively.

We next examine more fully how this relationship has changed over time by running a panel regression with time-specific coefficients

$$\log(Y_{i,t+k}) - \log(Y_{i,t}) = \alpha_t + \sum_{\tau=1850}^{1990} I\left(t=\tau\right)\left(\beta_\tau\eta_{i,t} + \lambda_\tau Y_{i,t}\right) + \epsilon_{i,t} \tag{29}$$

For the horizon $k = 20$ years.[13] Figure 9 plots the estimates of β_τ for each Census year along with the 90% confidence intervals based on standard errors clustered by occupation. The point estimates are negative for all but the 1860 and 1940 Censuses, and are significant in 1880, 1910, 1920, 1950, 1970, 1980, and 1990. The magnitude of this relationship has also stayed fairly constant over time. Thus occupations that are exposed to innovation have had measurably poor subsequent employment outcomes for a period of time spanning over 150

[11] We exclude farm occupations from the analysis.

[12] Census records for the year 1890 were destroyed in a fire, and so the employment growth observations for the 20-year horizon in 1870 or for the 10-year horizon in 1880 are not available.

[13] We focus on this horizon because we want to be able to include the 1880-1900 20-year period in our analysis.

years. It is worth noting that the 1930 quote in the introduction by John Maynard Keynes about technological unemployment appears especially prescient in the context of these findings; it occurred in the middle of a 20-year time period starting in 1920 which corresponded with a large innovation wave that was associated significant declines in employment for occupations whose tasks were related to those innovations.

We next split the Census years into innovation waves based on the periods identified by the Kelly et al. (2020) breakthrough innovation index to examine more closely how employment responds differentially during times of high innovation. Based on this we label the 20-year intervals starting in 1880, 1910, 1920, 1980, and 1990 as overlapping innovation waves, with the remaining years being non-innovation wave periods. Table 6 shows the estimates for the two subsamples. The point estimate is -.81 with a t-stat of -5.98 during innovation waves and -0.50 with a t-stat of -1.41 outside of innovation waves. Thus periods of greater leaps in the technological frontier are also more strongly associated with negative outcomes for exposed occupations, although the relationship is still negative in other years.

We also consider aggregations of the Census data at the occupation by industry level over time. We use the 1950 Census industry designations, which are available the furthest back in time. Because Census industry codes are unreliable before 1910 we start our analysis using the data from the 1910 Census. Let $Y_{i,j,t}$ represent the share of total non-farm employment for occupation i in industry j. We then estimate specifications of the following form for $k = 10$ and 20 years:

$$\log(Y_{i,j,t+k}) - \log(Y_{i,j,t}) = \alpha_0 + \beta\eta_{i,t} + \delta X_{t,j} + \epsilon_{i,j,t} \tag{30}$$

Here $X_{t,j}$ contain time, time and industry, or industry $\times$ time dummies depending on the specification, as well as the time t employment shares of the industry $\times$ occupation cell. The industry fixed effects pick up common employment shocks within an industry over the entire time period, while the industry $\times$ time fixed effects pick up time-specific within industry effects.

Table 7 presents the estimates of equation (30). Note that no matter the specification, the coefficients on $\eta_{i,t}$ are significantly negative. This means that the employment drops for high $\eta_{i,t}$ occupations are not merely due to overall declines in employment within industries that happen to employ workers with high technology exposure. Rather, much of the negative employment effects are found within rather than between industries. The fact that the coefficients do attenuate slightly when including industry fixed effects indicates that high $\eta_{i,t}$ occupations tend to be employed in industries that have experienced overall employment declines. Magnitudes in the 10-year specifications range from -.32% to -.67% at an annualized basis, while they range from -.6% to -.89% for the 20-year horizon. Thus a one-standard deviation increase in technological exposure can be expected to decrease employment for a given occupation within an industry by roughly 4-8% over the next 10 years and about 15-18% over the next 20 years.

As in our previous long-run analysis at the occupation× year level, these facts are also robust across different subsamples. In Table 8 we split the regressions into 1910-1950 and post-1960 time periods and focus on the 20-year horizon. In both subsamples $\eta_{i,t}$ predicts within-industry occupational employment declines. Interestingly, the magnitudes here appear to be a little higher in the earlier period subsample, indicating that the within-industry employment declines for exposed occupations are a more important part of the overall occupation-level job losses for the post-1960 time period than in the 1910-1950 time period.

3.2.2 Innovation exposure and growth in wages and employment post-1980

Having established the long-term relationship between new technologies and occupation-level outcomes, we now shift our focus more specifically on the post-1980 period spanning the information technology revolution. This period has much more data available so that we can examine shorter time intervals, as well as study the effects on wages. We start by looking at wages and employment data from the Current Population Survey Merged Outgoing Rotation Groups (MORG). We obtain the cleaned versions of MORG extracts provided by the Center for Economic Policy Research (CEPR).[14] Using these data we construct a time series of wage and employment growth for occupations at the occ1990dd level. Because occ1990dd cannot be crosswalked to a balanced panel of occupations using the Census 1970 occupation codes, we start our analysis in the post-1982 time period when these extracts began using the 1980 Census occupation classification scheme.

We estimate the following regressions

$$y_{i,t+k} - y_{i,t} = \alpha + \beta\eta_{i,t} + \delta X_{i,t} + \varepsilon_{i,t} \tag{31}$$

Here $y_{i,t}$ represents log wages or employment for a given occupation. The vector of controls $X_{i,t}$ includes three one-year lags of dependent variable, time fixed effects, wage, and the occupation employment share; we also include the Autor and Dorn (2013) routine-task intensity measure to test if $\eta_{i,t}$ survives the inclusion of this variable that has also been linked

[14] We use the "wage3" variable that combines the usual hourly earnings for hourly workers and non-hourly workers, which adjusts for top-coding using a lognormal imputation and is constructed to match the NBER's recommendation for the most consistent hourly wage series from 1979 to the present.

to technological unemployment. Because of the lags our first observation in the regression occurs in 1985. We examine horizons $k = 5, 10, 20,$ and 30 years. The dependent variable is again expressed in annualized percentage terms and $\eta_{i,t}$ is standardized.

The results are plotted in Figure 10, along with 90% confidence intervals. The responses for both wages and employment are strongly negative for all horizons. Employment decline magnitudes are similar to those we found in the long-run exercise using the IPUMS Census data, hovering around -0.6% on an annualized basis, while the decline in wages varies between 0.11% and 0.15% per year. These effects are statistically significant across all horizons, suggesting these are permanent effects.

In sum, we see that $\eta_{i,t}$ predicts negative movements in labor demand that reduce wages and employment, consistent with technological displacement of labor.

3.2.3 Which occupations are more exposed?

An informative descriptive fact in light of these findings concerns the average exposures by occupation skill levels. In Figure 11 we plot exposures against wage percentile ranks for the post-1980 period with the MORG wage data. Using wages as a proxy for worker skill, we find that the most exposed occupations tend to be found in the middle of the income distribution. Since $\eta_{i,t}$ predicts wage declines, and combined with the high exposure at the middle of the income distribution, our findings agree with the basic premise of Autor and Dorn (2013), who argue that middle skill occupations have experienced wage declines at the expense of low-skill service occupations. The two are especially consistent given that our measure is high for middle income occupations and that we find low exposures for service occupations that are specialize interpersonal communication.

3.2.4 Occupations vs Industry

Our analysis so far does not differentiate among workers of the same occupation that work in different industries. We may expect however that a given occupation may have a different meaning across different industries. Further, because many of the most exposed occupations come from manufacturing type occupations one may worry that industry specific shocks can explain our findings. Here, we aim to differentiate between industry and occupations for the 1980-2012 time period.

We estimate a cross-sectional regression of long-run differences using the combined information on wages and employment in the 1980 Census and the 2012 ACS. In particular, we use the 1980 Census and 2012 ACS data from Deming (2017), which are reported at the occupation by industry by education level and aggregate to data to industry by occupation. We then estimate a long-difference cross-sectional specification similar to Webb (2019) as follows

$$\Delta y_{i,j} = \alpha + \alpha_j + \beta \eta_{i,1980} + \delta X_i + \epsilon_{i,j} \tag{32}$$

Here i indexes occupations and j indexes industries. The outcome $\Delta y_{i,j}$ denotes either the log change in employment or the change in log wages over the 1980-2012 time period. We include industry fixed effects α_j to account for industry specific shocks that may be correlated with occupational outcomes. Controls X_i include occupation employment share in 1980, occupation log wage in 1980, three indicators for the occupations education level in 1980, the routine-task intensity and the measure of occupation-level offshorability from Autor and Dorn (2013), and the Webb (2019) measures of exposure to robotics or software patents, depending on the specification. We weight observations by the employment share in 1980 and cluster standard errors by industry. Tables 9 and 10 report the results from estimating (32) for employment and wages, respectively.

The first row of Table 9 shows that our results are not driven by industry-specific shocks. In particular, the estimated coefficient β is negative and statistically significant even though industry-fixed effects are included in equation (32). This result implies that differences in the dependent variable $\eta_{i,1980}$ are related to differences in subsequent employment and wage growth across occupations in the same industry. Moreover, including the Autor and Dorn (2013) or Webb (2019) measures has little impact on the coefficients, nor the significance, indicating that $\eta_{i,t}$ contains independent information relative to these alternate metrics.

3.2.5 Innovation exposure, employment, and recessions

We next examine how employment in exposed occupations responds around recessions. Jaimovich and Siu (2018) document that "routine" jobs explain a large portion of employment declines around the last three recessions. However, they assign very broad groups of occupations into the routine category, while there is heterogeneity in measures of routineness for the specific occupations within these groups. We show that a distinct pattern of job losses emerges around recessions by looking at occupations which are in the top quintile of $\eta_{i,t}$. We again use the CPS MORG data and start in 1985, which is the year our regressions in Figure 10 began. In Figure 12 we see that occupation which were in the top quintile of $\eta_{i,t}$ in 1985 experienced stark declines in employment around the 1991, 2001, and 2007-2008 recessions, with a flatter but slightly declining profile in between recessions. Meanwhile, assigning occupations into the top quintile of routine-task intensity in 1985, we do see a persistent decline over the time period, but a much less pronounced pattern around recessions.

This pattern agrees with models of innovation-related job displacement where the opportunity to replace labor with an automation technology is a real option for firms. For example, Zhang (2019) shows that in a production based asset pricing model where firms choose to invest in labor-saving technologies, they choose to exercise this option when expected cash flows are temporarily low. Therefore the pattern exhibited in Figure 12 is consistent with employers replacing high $\eta_{i,t}$ workers with capital when the exercise value for doing so is high.

3.3 Industry-Level Productivity and the Labor Share

A defining feature of labor-displacing technological change is the divergence in subsequent outcomes between industries (or firms) that implement these technologies and the subset of the industries' labor force that are most exposed to them. On the one hand, by reducing the costs of production labor-saving technologies can be expected to improve productivity and increase output. On the other hand, technologies that substitute for tasks performed by labor also can be expected to have the effect of reducing the labor share of affected occupations. This is in contrast to a Hicks-neutral productivity shift, which increases output and productivity but leaves factor shares unchanged. We use our measures of occupation task by patent textual similarity and the breakthrough patents from Kelly et al. (2020) to construct an index of labor-displacing technological change. We then use the NBER manufacturing database to examine the effects on industry performance and the labor force within an industry.

More specifically, first consider our adjusted measure of patent by occupation task textual similarity, $\widetilde{\rho}_{i,p}$ from Section 1, where i indexes occupations and p is now used to index patents (with the sample restricted here to the set of breakthrough patents as defined in (12)). Denote the industry j employment share of occupation i at time t with $w_{i,j,t}$.[15] For breakthrough patent p issued in year t and for manufacturing industry j, define the similarity of patent p and industry j by

$$\theta_{j,p} = \sum_i w_{i,j,t}\widetilde{\rho}_{i,p} \tag{33}$$

Next, let q^{θ}_{p50} denote the median of $\theta_{j,p}$ and let κ_t represent the US population in time t

[15] We crosswalk occupations to occ1990dd codes from Autor and Dorn (2013), and our DOT text data is defined at the 2010 SOC code level, so $\widetilde{\rho}_{i,p}$ is technically the average patent × occupation adjusted similarity within an occ1990dd code. We use IPUMS Census data from 1960-2010 based on IPUMS ind1990 industry codes get occupation-industry employment shares. To obtain employment shares on non-Census years we linearly interpolate between IPUMS occupation-industry employment shares for the nearest prior and future decades.

as before. For each breakthrough patent p we assign it to an industry using probabilistic patent CPC tech class to NAICS crosswalks constructed by Goldschlag, Lybbert, and Zolas (2020).[16] Label the set of breakthrough patents issued in year t by Γ_t, and $\alpha_{j,p}$ the probability of breakthrough patent p being issued to industry j. We then define the industry-level breakthrough patent index (including only patents with high average textual similarity to the industry workforce) by

$$\psi_{j,t} = \frac{1}{\kappa_t} \sum_{p \in \Gamma_t} \alpha_{j,p} \mathbf{1}(\theta_{j,p} \geq q^{\theta}_{p50}) \tag{34}$$

Having constructed $\psi_{j,t}$, we use it to predict industry outcomes using the NBER manufacturing database. In particular, we run the following regression for various outcome variables X:

$$\log(X_{j,t+k}) - \log(X_{j,t}) = \alpha + \beta\psi_{j,t} + \delta Z_{j,t} + \epsilon_{j,t} \tag{35}$$

The horizon k is set to 20 years. Controls include the log level and the lagged 5-year growth rate of the outcome variable, industry employment shares, and year fixed effects. We display the results in Table 11, scaling $\psi_{j,t}$ to unit standard deviation and expressing growth rates of the dependent variable in annualized percentage terms.

In panel A we look at the following industry outcomes: total factor productivity, investment (capital expenditures), value-added, and value-added per worker. Consistent with a positive technology shock, the measure significantly predicts each of these outcomes with positive sign. Next, in panel B we examine the subsequent growth in the labor share (defined as wages over value added); the labor share of non-production workers; the labor share of production workers; and the share of wages accruing to production workers. The evidence suggests that $\psi_{j,t}$ captures labor-saving technological change concentrated on the tasks of a specific subset of the labor force: production workers. In particular, while $\psi_{j,t}$ does predict overall declines in the labor share, this is driven entirely by production workers. A one-standard deviation increase in $\psi_{j,t}$ decreases the share of value added accruing to production workers' wages by about 0.6% per year; the share of industry wages paid to production workers also declines by 0.3% per year in response to a one-standard deviation increase in $\psi_{j,t}$. These findings support our previous occupation-level analysis, as production-type occupations are among the most highly exposed and the most negatively affected.

Overall, the empirical results presented in this section paint a clear picture of technological

[16] Since our Census data is at the ind1990 level, we then use NAICS to SIC and SIC to ind1990 crosswalks from Autor, Dorn, and Hanson (2013) to get a probability weight for each patent at the Census ind1990 level.

displacement of labor. Our measure consistently predicts a decline in employment across the last 150 years; similarly, our measure also predicts declines in wages in the post-1980 subsample where more detailed occupation-level wage data are available. Finally, an industry level index of labor-saving technologies constructed from our patent-by-occupation textual similarity scores predicts divergent outcomes for innovating industries and the production workers employed in these industries. Though these estimates are correlational rather than causal, they are strongly in line with theories linking labor-saving innovations to negative demand shocks for specific types of labor.

3.4 Alternate Approaches

The evidence presented so far suggests that our occupation exposure measure mostly identifies labor-saving technological innovations. This feature is partly embedded in its construction: our measure of occupational technological exposure yields a high exposure score for occupations whose task descriptions look similar to the description of patents. As labor-saving technologies are introduced, they will execute the same tasks that workers once performed.

As a concrete example, consider US patent number 6,289,319, titled “Automatic business and financial transaction processing system”, and which as shown in Table 1 is the most similar patent to the “Loan Interviewers and Clerks” occupation. The DOT task description indicates that a person with this occupation “calls or writes to credit bureaus, employers, and personal references to check credit and personal references.” Meanwhile, the description of this patent says that “Loan processing has traditionally been a labor-intensive business...the principal object of this invention is to provide an economical means for screening loan applications.” This appears to be an example of a technology which has high potential to be labor saving because it is intended to do the same tasks performed manually by a worker in a more efficient manner.

However, there can also be cases where new technologies improve the productivity of tasks that workers were already doing. Our measure may occasionally pick up such instances. Consider an example from the “Database Administrators” occupation (SOC code 151141). According to the DOT, a database administrator “coordinates physical changes to computer databases.” One of the most similar patents to this occupation is US patent number 5,093,782, entitled “Real time event driven database management system.” This patent indicates that it provides “a database management system which is capable of supporting processes requiring the updating and retrieval of data elements at a high rate.” This is likely to make the work of database administrators more efficient and hence looks more likely to be labor productivity-enhancing for this occupation.

Given our methodology, it is not obvious ex-ante whether our measure will primarily capture technologies that enhance the productivity of existing labor, or whether it captures technologies that displace the human capital of incumbent workers. If our measure captures both aspects of innovation, interpreting the effects in the previous section is somewhat challenging, as we are averaging across labor-saving versus labor enhancing innovations.

We next undertake a statistical prediction exercise to examine how much our measure may be diluted by mixing labor-saving and productivity-enhancing innovations. Specifically, we leverage recent advances in topic modeling to construct a composite predictor from patent text whose purpose is to maximize the *in-sample* predictability of employment declines. This measure is akin to a principal component; it has no straightforward economic interpretation, but it rather provides a benchmark for how much greater the displacive effects we uncover could possible be.

To do so, we use a method proposed by Cong, Liang, and Zhang (2019), which is well-suited to prediction exercises using large-scale textual data. We focus on the period of time covered by our CPS merged outgoing rotation group sample (1985-2018) used in the employment regressions in Figure 10—which is the period where our employment and wage data coverage is most comprehensive. We next briefly describe the steps of how we adapt their procedure; Section 5.2 describes the procedure in more detail.

The objective of the procedure is to form a group of text-based predictors (textual factors) which can be used to predict employment outcomes. The starting point is the vector representations of word meanings discussed in Section 1. Using the vector representations of word meanings we implement a clustering routine to form groups of related words as a guess for the topics which are relevant in the set of patents. This yields about 140,000 word clusters, or candidate "topics". We then follow Cong et al. (2019) in choosing only the subset of candidate topics that are most relevant in the actual patent documents, which will be those that arise most frequently in the patent text. Using their ranking criteria we choose the top 500 most important topics in the set of breakthrough patents from Kelly et al. (2020). Next we compute the loadings on each topic (a document-level measure of topic importance) for every breakthrough patent document and also every occupation. Finally, multiplying the sum of all patent loadings on a given topic in a given year times an occupation's topic loading yields a time-varying measure of occupation exposure to that topic. Doing this we obtain each occupation's exposure in each year to the 500 topics (textual factors) found to be the most important in the set of breakthrough patents. In Appendix Figure A.1 we show four of the topics generated by the LSH routine that were identified as important using the Cong et al. (2019) ranking metric. These topics include words related roughly to engines and energy; mechanical parts and tools; electronic communication; and computer systems.

We next use these 500 textual factors to form a single predictor that is optimized to predict occupation declines in-sample. We design a simple exercise to do so. First, we regress the 10-year employment growth rate for occupations in the CPS merged outgoing rotation group sample on each of the individual candidate predictors $\psi_{i,k,t}$ in univariate regressions, also including a time fixed effect in each regression.[17] We then retain each of the individual predictors which have a t-stat below -1.96, where standard errors have been clustered at the occupation level. A total of 185 out of the 500 topics satisfy the cutoff for negatively predicting employment growth, considerably more than would be expected by random chance. In order to quantify how much information in the patent text can be leveraged to predict occupation expansions, we also perform the analogous exercise for positive predictors of employment growth; this yields just 17 significantly positive predictors, less than would even be expected by random chance.

For each of the sets of extracted factors (coming from either the group identified as negative or positive predictors of employment growth) we create a combined predictor by taking the cross-sectional mean across all individual predictors or by taking the first principal component.[18] We label the variable constructed to negatively predict employment $\xi_{i,t}$, in line with the labor-saving technology parameter in the model from section 3.1, and the version intended to positively predict employment as $\gamma_{i,t}$ to represent the labor productivity enhancing parameter. Appendix table A.1 demonstrates unsurprisingly that both measures strongly predict employment with the correct sign in-sample. In addition we include an out-of-sample test based on wage growth, and find that they also strongly predict wage growth with the same sign despite the fact that wage growth is not highly correlated with employment growth overall. However, our constructed variable $\xi_{i,t}$, meant to predict employment declines, has much stronger in-sample performance for both employment and wage growth than the alternative measure $\xi_{i,t}$ meant to predict positive employment growth.

The final step is to examine how our original measure $\eta_{i,t}$ correlates with these two statistical predictors of employment. Despite the fact that these measures are estimated using different approaches, in appendix table A.2 we show that the correlation between our baseline measure $\eta_{i,t}$ and the statistical predictor constructed to represent exposure to labor-saving technologies $\xi_{i,t}$ is approximately 74 percent. If we further exclude information/communication technology and electronics patents in computing $\eta_{i,t}$ this correlation

[17]Occupations are again collapsed to the occ1990dd-level by taking averages of 6-digit 2010 SOC codes within an occ1990dd occupation. We also remove from the analysis 9 of the 500 predictors that had non-zero loadings for either no occupations or just a single occupation.

[18]The two turn out to be almost exactly the same, with correlations around .99.

increases to 0.92. Meanwhile, the version of $\eta_{i,t}$ that only includes information/communication technology and electronics patents has a much lower correlation with $\xi_{i,t}$, and is the only version with a statistically significantly positive correlation of 8.7 percent with $\gamma_{i,t}$, our constructed predictor which is intended to positively predict employment.

This exercise demonstrates that constructing a variable which is explicitly data mined to predict employment declines in sample yields an object that highly resembles our occupational technological exposure measure $\eta_{i,t}$. Further, it appears harder to data mine a predictor which positively predicts employment, as there are far fewer candidate topics that do so. Therefore, as intended our measure $\eta_{i,t}$ is more successful at capturing exposure to technologies that substitute for, rather than complement, human capital. That being said, because of the rise of ICT patents in the end of the 20th century these two channels are likely mixed somewhat towards the end of the sample. In unreported regressions we find that excluding ICT patents yields a stronger predictor of employment and wage declines in the CPS merged outgoing rotation group sample. It's likely that this is much less of a factor in the early part of our sample; as documented in section 2.2, ICT patents only started to change the skill composition of innovation exposure towards cognitive tasks around the 1980s.

4 Conclusion

In this paper we contribute to the literature on technological innovation and labor displacement by introducing a method for measuring the similarity between occupation task descriptions and new technologies as represented in the text of patents. We use the resulting measure to examine the type of occupational tasks which are exposed to innovation, demonstrating that while non-routine manual (physical) and routine-manual tasks have been highly exposed throughout the last 150+ years, the innovations of the information technology revolution in the post-1980 period saw an increased relationship with cognitive tasks. We also find that our measure is strongly correlated with employment declines since the late 19th century, as well as drops in both wages and employment in the post-1980 period. This provides evidence directly linking occupational employment and wage declines over a long period of time to new technologies which are likely to substitute for labor, suggesting that technological displacement of labor has been a persistent phenomenon over the past century and a half.

References

Acemoglu, D. and D. Autor (2011). Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of Labor Economics, Volume 4. Amsterdam: Elsevier-North*, pp. 1043–1171.

Acemoglu, D. and P. Restrepo (2018, June). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review 108*(6), 1488–1542.

Acemoglu, D. and P. Restrepo (2020). Robots and jobs: Evidence from us labor markets. *Journal of Political Economy forthcoming*.

Arora, S., Y. Liang, and T. Ma (2017). A simple but tough-to-beat baseline for sentence embeddings. In *ICLR*.

Atack, J., R. A. Margo, and P. W. Rhode (2019, May). "automation" of manufacturing in the late nineteenth century: The hand and machine labor study. *Journal of Economic Perspectives 33*(2), 51–70.

Autor, D. H. and D. Dorn (2013, August). The growth of low-skill service jobs and the polarization of the us labor market. *American Economic Review 103*(5), 1553–97.

Autor, D. H., D. Dorn, and G. H. Hanson (2013, October). The china syndrome: Local labor market effects of import competition in the united states. *American Economic Review 103*(6), 2121–68.

Autor, D. H., L. F. Katz, and M. S. Kearney (2006, May). The Polarization of the U.S. Labor Market. *American Economic Review 96*(2), 189–194.

Autor, D. H., F. Levy, and R. J. Murnane (2003, 11). The Skill Content of Recent Technological Change: An Empirical Exploration. *The Quarterly Journal of Economics 118*(4), 1279–1333.

Cong, W., T. Liang, and X. Zhang (2019). Textual factors: A scalable, interpretable, and data-driven approach to analyzing unstructured information.

Deming, D. J. (2017, 06). The Growing Importance of Social Skills in the Labor Market. *The Quarterly Journal of Economics 132*(4), 1593–1640.

Goldin, C. and L. Katz (2008). *The Race Between Education and Technology*. Harvard University Press.

Goldin, C. and L. F. Katz (1998, August). The origins of technology-skill complementarity. *The Quarterly Journal of Economics 113*(3), 693–732.

Goldschlag, N., T. J. Lybbert, and N. J. Zolas (2020). Tracking the technological composition of industries with algorithmic patent concordances. *Economics of Innovation and New Technology 29*(6), 582–602.

Goos, M. and A. Manning (2007). Lousy and lovely jobs: The rising polarization of work in britain. *The Review of Economics and Statistics 89*(1), 118–133.

Hall, B. H., A. Jaffe, and M. Trajtenberg (2005). Market value and patent citations. *The RAND Journal of Economics 36*(1), 16–38.

Hornstein, A., P. Krusell, and G. L. Violante (2005, Fall). The Effects of Technical Change on Labor Market Inequalities. In P. Aghion and S. Durlauf (Eds.), *Handbook of Economic Growth*, Volume 1 of *Handbook of Economic Growth*, Chapter 20, pp. 1275–1370. Elsevier.

Hornstein, A., P. Krusell, and G. L. Violante (2007). Technology—Policy Interaction in Frictional Labour-Markets. *Review of Economic Studies 74*(4), 1089–1124.

Jaimovich, N. and H. Siu (2018). The trend is the cycle: Job polarization and jobless recoveries. *Review of Economics and Statistics, forthcoming*.

Jerome, H. (1934). *Mechanization in Industry*. NBER.

Kelly, B., D. Papanikolaou, A. Seru, and M. Taddy (2020). Measuring technological innovation over the long run. *American Economic Review: Insights forthcoming*.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman (2017, 03). Technological Innovation, Resource Allocation, and Growth. *The Quarterly Journal of Economics 132*(2), 665–712.

Mikolov, T., I. Sutskever, K. Chen, G. Corrado, and J. Dean (2013). Distributed representations of words and phrases and their compositionality. *CoRR abs/1310.4546*.

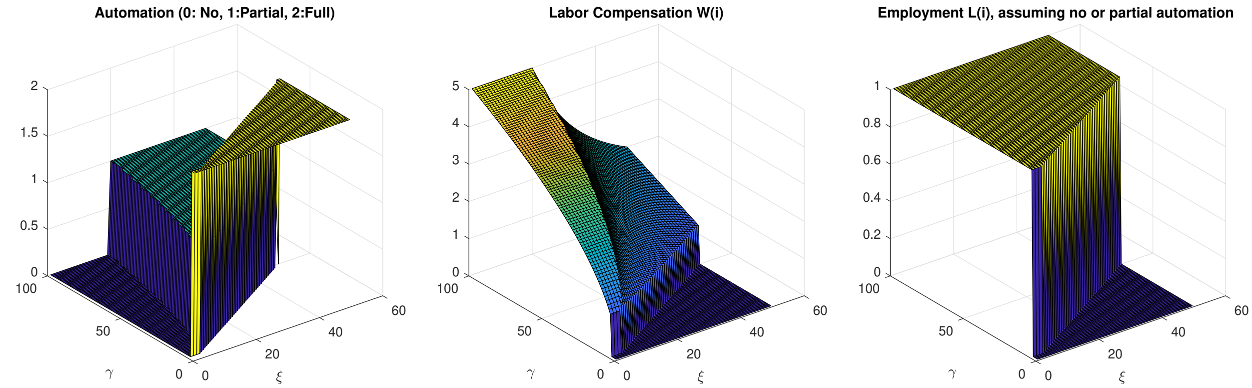
Pennington, J., R. Socher, and C. D. Manning (2014). Glove: Global vectors for word representation. In *EMNLP*.

Webb, M. (2019). The impact of artificial intelligence on the labor market.

Zhang, M. B. (2019). Labor-technology substitution: Implications for asset pricing. *The Journal of Finance 74*(4), 1793–1839.

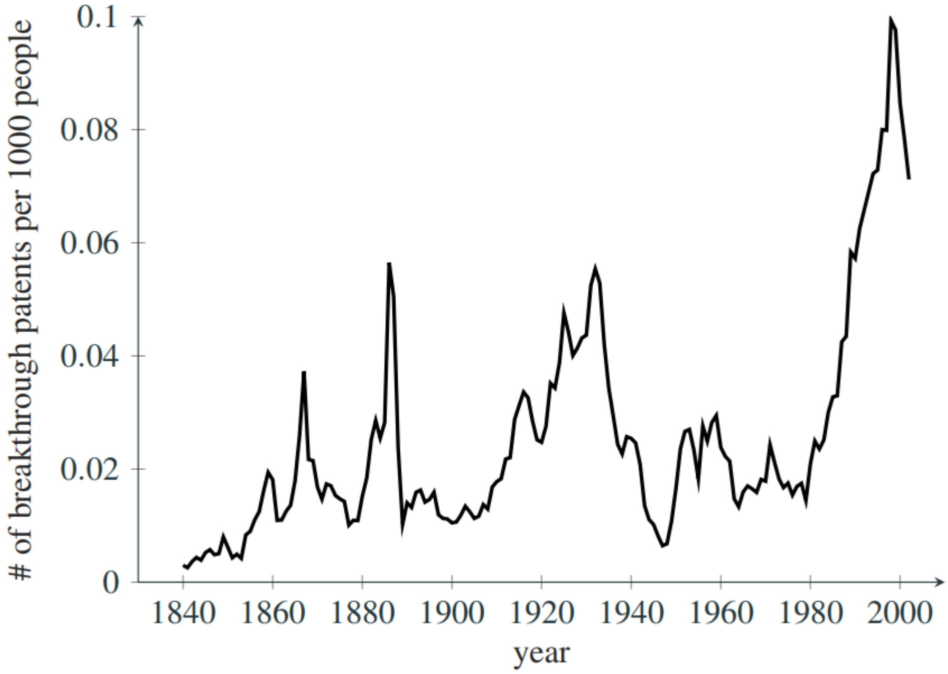
Figures and Tables

Figure 1: Technology and Worker Outcomes



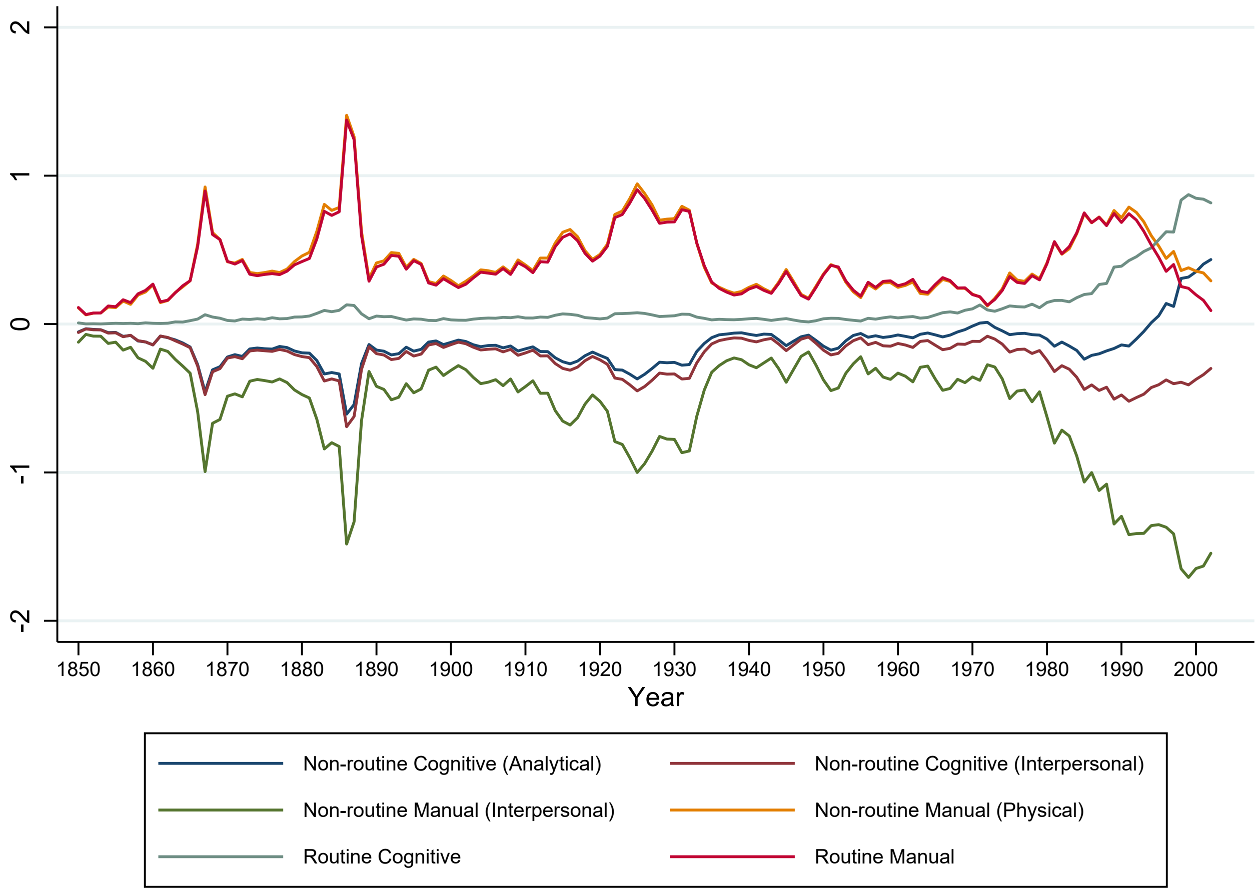
This figure plots model-implied values for the propensity to automate, wages, and employment of tasks as a function of labor productivity in the task, $\gamma(i)$, and capital productivity in the task, $\xi(i)$. See the model in section 3.1 for details.

Figure 2: Breakthrough Patents Over Time



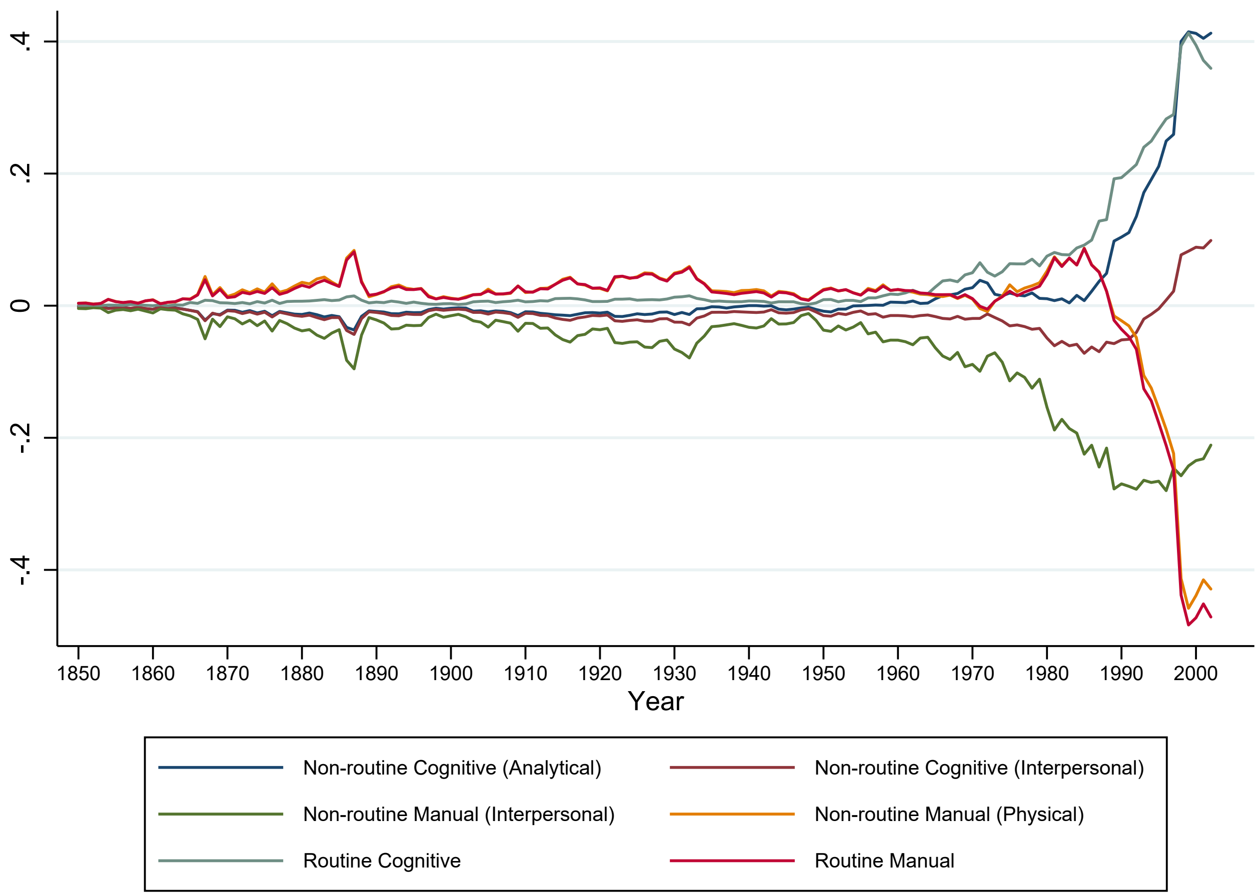
This figure plots the Kelly et al. (2020) index of breakthrough patents scaled by US population (in thousands).

Figure 3: Exposures of Tasks To Technological Change–All Patents



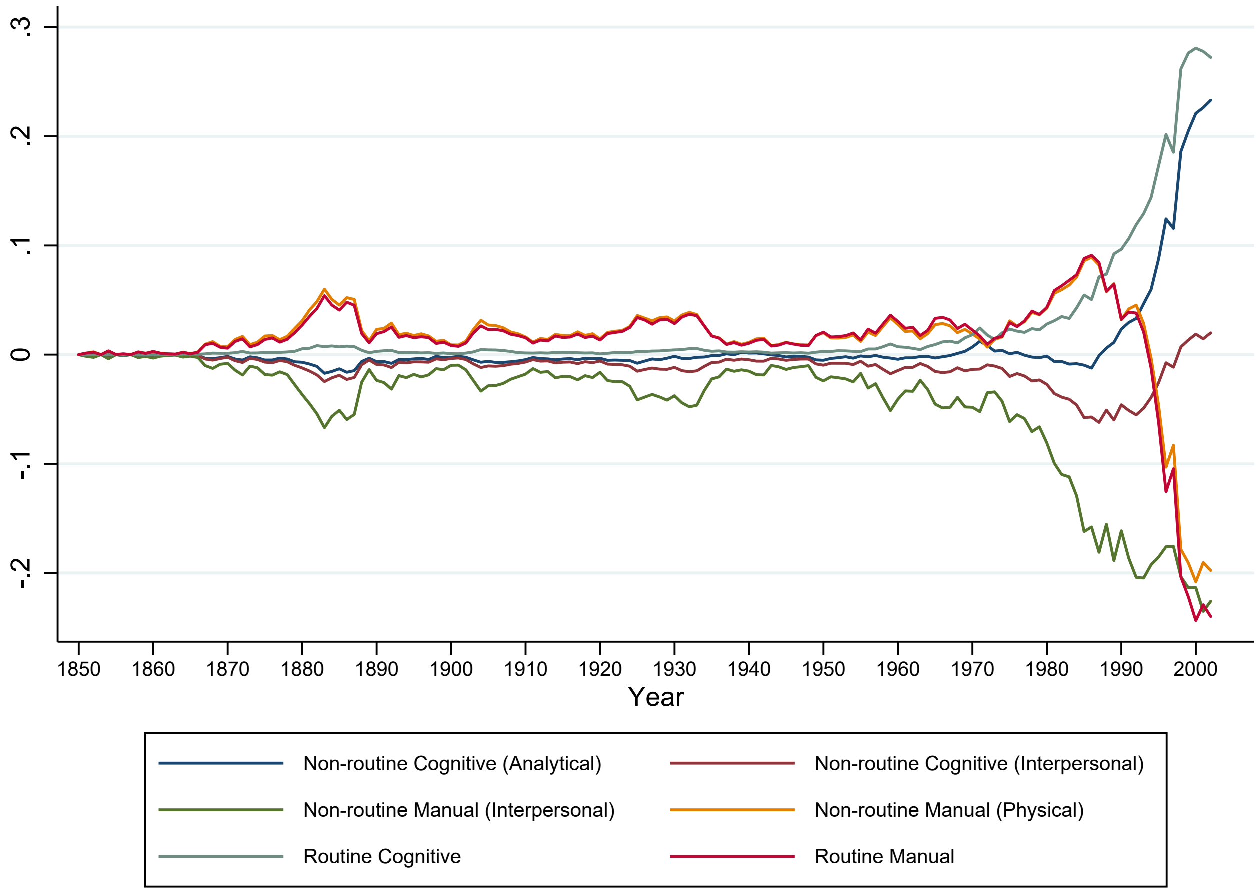
This figure plots the time series of $\lambda_{w,t}$, the task-level exposures to technology as described in the main text. We use the occupation task types defined in Acemoglu and Autor (2011). The plots include all patents in a given year. See section 2.2 in the text for further details.

Figure 4: Exposures of Tasks To Technological Change–Information & Communications Patents



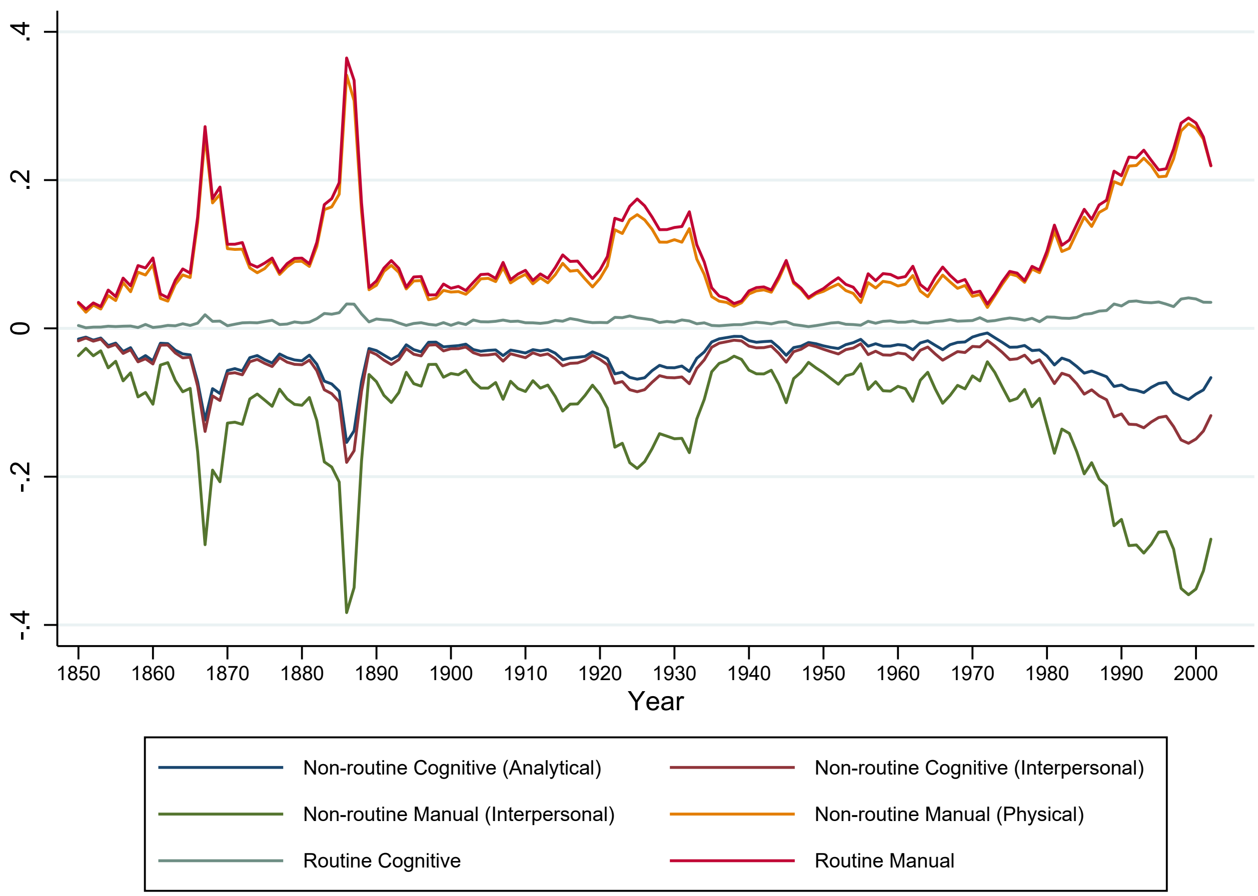
This figure plots the time series of $\lambda_{w,t}$, the task-level exposures to technology as described in the main text. We use the occupation task types defined in Acemoglu and Autor (2011). The plots include patents assigned to the information and communications broad patent technology subclass, as defined in Kelly et al. (2020). See section 2.2 in the text for further details.

Figure 5: Exposures of Tasks To Technological Change–Electronics Patents



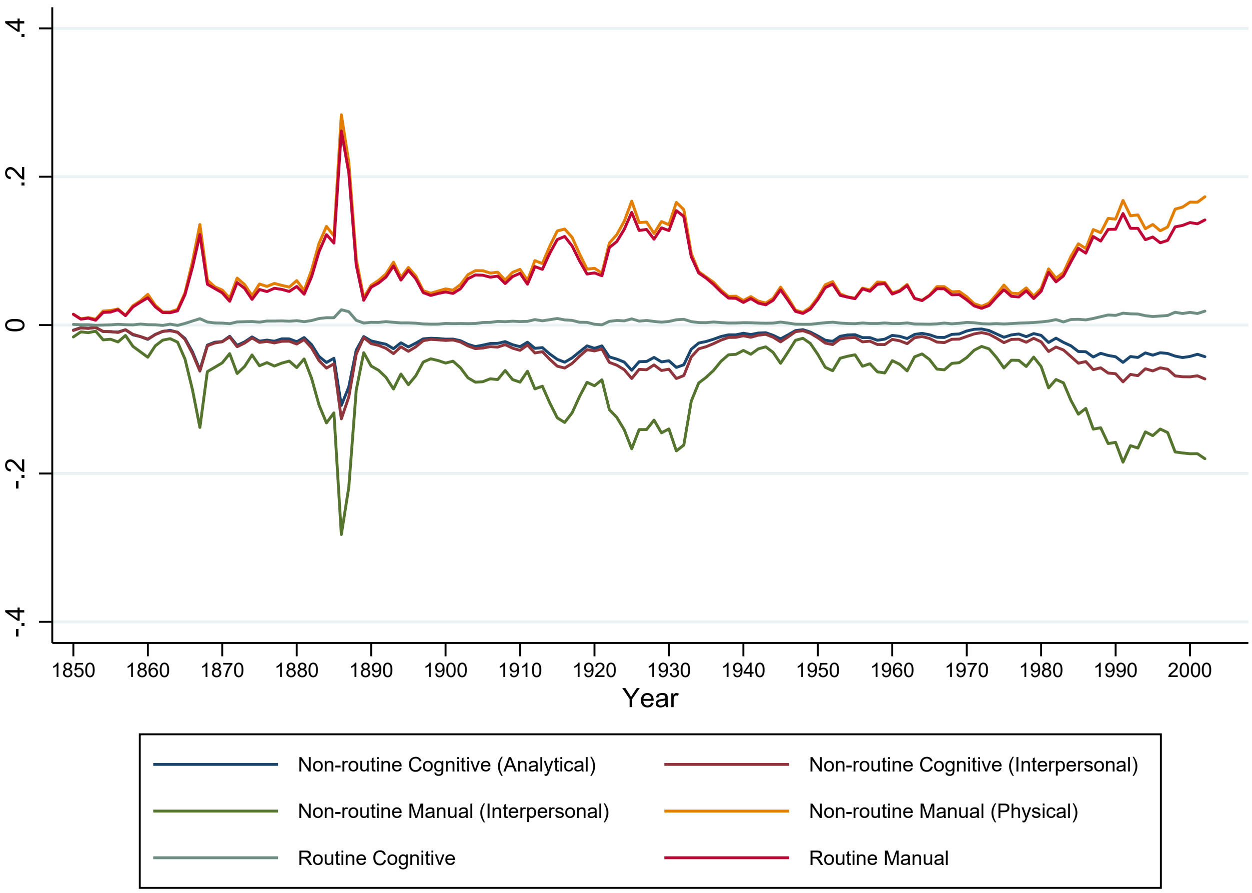
This figure plots the time series of $\lambda_{w,t}$, the task-level exposures to technology as described in the main text. We use the occupation task types defined in Acemoglu and Autor (2011). The plots include patents assigned to the electronics broad patent technology subclass, as defined in Kelly et al. (2020). See section 2.2 in the text for further details.

Figure 6: Exposures of Tasks To Technological Change–Manufacturing Process Patents



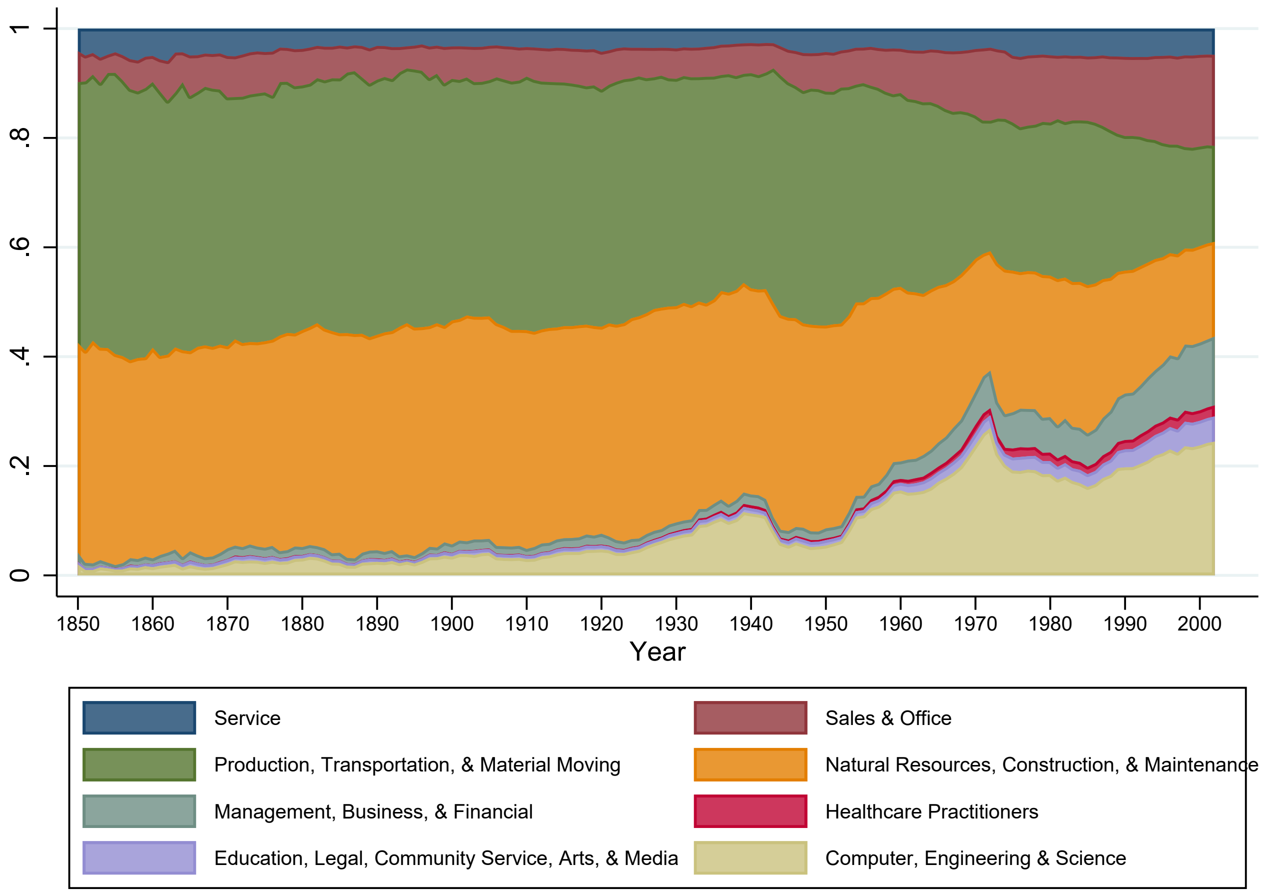
This figure plots the time series of $\lambda_{w,t}$, the task-level exposures to technology as described in the main text. We use the occupation task types defined in Acemoglu and Autor (2011). The plots include patents assigned to the manufacturing process broad patent technology subclass, as defined in Kelly et al. (2020). See section 2.2 in the text for further details.

Figure 7: Exposures of Tasks To Technological Change–Construction/Engineering Patents



This figure plots the time series of $\lambda_{w,t}$, the task-level exposures to technology as described in the main text. We use the occupation task types defined in Acemoglu and Autor (2011). The plots include patents assigned to the construction and engineering broad patent technology subclass, as defined in Kelly et al. (2020). See section 2.2 in the text for further details.

Figure 8: Technological Change And Occupations: Composition



This figure plots the average of our occupation-level innovation exposure index, $\eta_{i,t}$, where $\eta_{i,t}$ has been averaged separately within eight broad occupation groups. The occupation group averages are re-scaled each year so that the total across all groups sums to one in the given year.

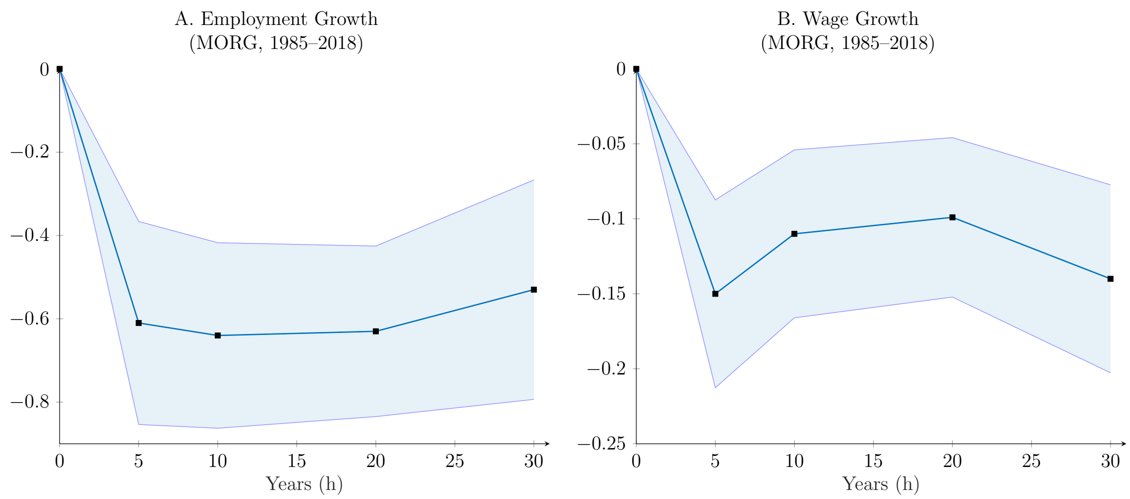
Figure 9: Yearly Regressions of 20-Year Employment Share Growth on $\eta_{i,t}$ for Census Years 1850-1990

Plot above shows coefficients from regression of form

$$\log(Y_{i,t+k}) - \log(Y_{i,t}) = \alpha_t + \sum_{\tau=1850}^{1990} I\left(t=\tau\right)\left(\beta_\tau \eta_{i,t} + \lambda_\tau Y_{i,t}\right) + \epsilon_{i,t}$$

Here $Y_{i,t}$ is the occupation's share in total non-farm employment. Standard errors are clustered by occupation and shaded area represents the corresponding 90% confidence intervals for β_τ. Growth rates are expressed in annualized percentage terms and $\eta_{i,t}$ is standardized.

Figure 10: Employment and Wage Responses to $\eta_{i,t}$

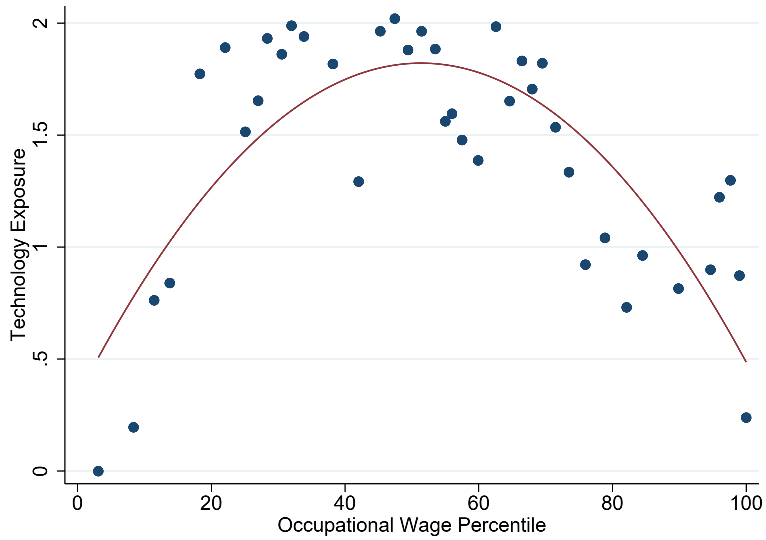


The Figures above plot coefficients from panel regressions of annualized wage and income growth rates over different time horizons on occupation innovation exposures:

$$y_{i,t+k} - y_{i,t} = \alpha + \beta\eta_{i,t} + \delta X_{i,t} + \varepsilon_{i,t}$$

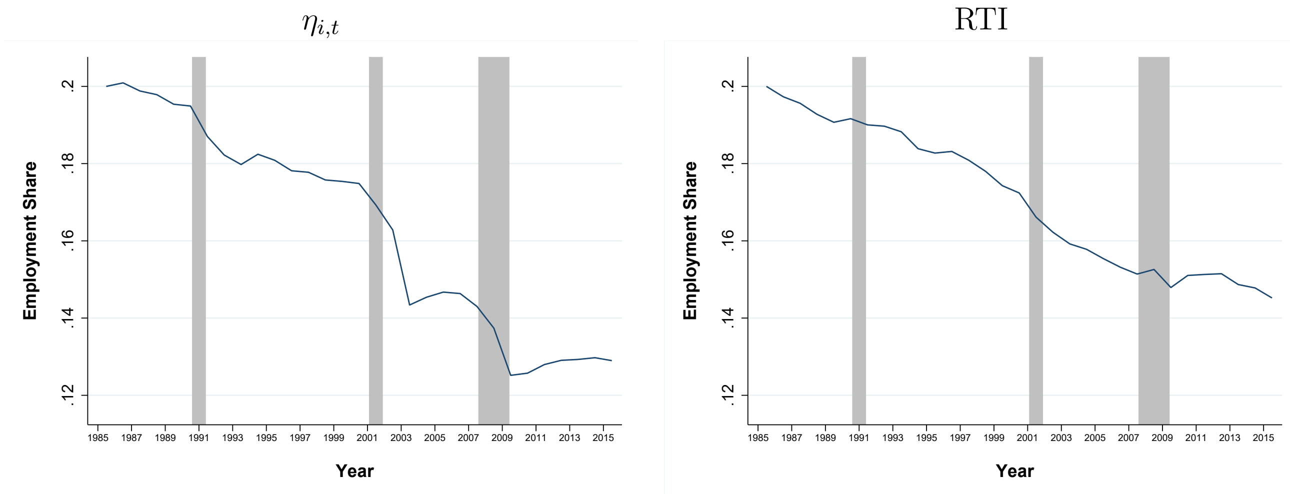
Controls $X_{i,t}$–includes three one-year lags of dependent variable, time fixed effects, wage, and occupation employment share. Dependent variable is expressed in annualized percentage terms and $\eta_{i,t}$ is standardized. Figures plot 90% confidence interval for each time horizon. Data come from the CPS Merged Outgoing Rotation Groups (MORG).

Figure 11: Occupational Innovation Exposures By Income Percentile Rank



This figure plots average $\eta_{i,t}$ for occupations over the 1980 to 2002 period by wage percentile rank. The wage data come from the Current Population Survey Merged Outgoing Rotation Groups.

Figure 12: Employment Share Over Time for Occupations In Top $\eta_{i,t}$ and Routine Task Intensity (RTI) Quintiles in 1985



The above figure plots aggregate employment shares over time for occupations that were in the top quintiles of innovation exposure ($\eta_{i,t}$) and routine-task intensity in the year 1985. Vertical shaded bars represent NBER recession dates. Data source: CPS Merged Outgoing Rotation Groups extracts obtained from the Center For Economic Policy Research website.

Table 1: Most Similar Patents For Select Occupations

Cashiers (SOC Code 412011)	
5055657	Vending type machine dispensing a redeemable credit voucher upon payment interrupt
5987439	Automated banking system for making change on a card or user account
5897625	Automated document cashing system
6012048	Automated banking system for dispensing money orders, wire transfer and bill payment
5598332	Cash register capable of temporary-closing operation
Loan Interviewers and Clerks (SOC Code 434131)	
6289319	Automatic business and financial transaction processing system
5611052	Lender direct credit evaluation and loan processing system
6233566	System, method and computer program product for online financial products trading
5940811	Closed loop financial transaction method and apparatus
5966700	Management system for risk sharing of mortgage pools
Railroad Conductors (SOC Code 534031)	
5828979	Automatic train control system and method
6250590	Mobile train steering
3944986	Vehicle movement control system for railroad terminals
6135396	System and method for automatic train operation
5797330	Mass transit system

Table 2: Most Similar Occupations For Select Patents

"Knitting-machine" (Patent No. 276146, Issued in 1883)
Textile Knitting and Weaving Machine Setters, Operators, and Tenders
Sewing Machine Operators
Sewers, Hand
Fabric Menders, Except Garment
Textile Winding, Twisting, and Drawing Out Machine Setters, Operators, and Tenders
"Metal wheel for vehicles" (Patent No. 1405358, Issued in 1922)
Automotive Service Technicians and Mechanics
Cutting, Punching, and Press Machine Setters, Operators, and Tenders, Metal and Plastic
Maintenance Workers, Machinery
Grinding, Lapping, Polishing, and Buffing Machine Tool Setters, Operators, and Tenders, Metal and Plastic
Rolling Machine Setters, Operators, and Tenders, Metal and Plastic
"System for managing financial accounts by a priority allocation of funds among accounts" (Patent No. 5911135, Issued in 1999)
Financial Managers
Credit Analysts
Loan Interviewers and Clerks
Accountants and Auditors
Bookkeeping, Accounting, and Auditing Clerks

Table 3: Occupations Most and Least Exposed to Innovation

Top 5 Occupations by Average $\eta_{i,t}$	Bottom 5 Occupations by Average $\eta_{i,t}$
Inspectors, Testers, Sorters, Samplers, and Weighers	Mental Health Counselors
Metal Workers and Plastic Workers, All Other	Dancers
Cutting, Punching, and Press Machine Setters, Operators, and Tenders, Metal and Plastic	Funeral Attendants
Production Workers, All Other	Judges, Magistrate Judges, and Magistrates
Electromechanical Equipment Assemblers	Clergy

Table 4: Unconditional Correlations of $\eta_{i,t}$ With Task Categories

NR Cog (Analytical)	-0.12**					
	(-2.53)					
NR Cog (Interpersonal)		-0.16***				
		(-4.65)				
NR Man (Physical)			0.24***			
			(5.65)			
NR Man (Interpersonal)				-0.33***		
				(-8.43)		
Routine Cognitive					0.033	
					(0.95)	
Routine Manual						0.24***
						(5.74)

This figure plots the correlations of $\eta_{i,t}$ with the occupation task types computed from O*NET in Acemoglu and Autor (2011). Correlations are weighted by the Acemoglu and Autor (2011) occupation employment weights used to normalize the distribution of tasks to mean zero and standard deviation one.

Table 5: Technology And Employment Over the Long Run (1850-2010)

Full Sample			Subsamples (20 Year Horizon)		
	10 Years	20 Years		1850-1920	1930-1990
$\eta_{i,t}$	-0.41***	-0.70***	$\eta_{i,t}$	-0.94***	-0.65***
	(-4.18)	(-4.59)		(-2.66)	(-4.34)
Time FE	X	X	Time FE	X	X
Emp Shares	X	X	Emp Shares	X	X
N	2876	2585	N	978	1607
R^2 (Within)	0.015	0.040	R^2 (Within)	0.034	0.080

The table above plots results from regressions of the form

$$\log(Y_{i,t+k}) - \log(Y_{i,t}) = \alpha_0 + \alpha_t + \beta\eta_{i,t} + \lambda Y_{i,t} + \epsilon_{i,t}$$

for $k = 10, 20$ years for Census years spanning from 1850-2010. Here $Y_{i,t}$ is the occupation's share in total non-farm employment. $\eta_{i,t}$ is standardized and growth rates are in annualized percentage terms. Standard errors are clustered by occupation and corresponding t-stats are shown in parentheses. Observations are weighted by occupation employment share at time t. Census year 1870 does not show up in the first column of the 20-year subsample regressions because the 1890 Census records no longer exist.

Table 6: Technology And Employment During and Outside of Innovation Waves, 20-year Horizon Employment Growth

	Innovation Wave	Other Years
$\eta_{i,t}$	-0.81***	-0.50
	(-5.98)	(-1.41)
Time FE	X	X
N	1109	1476
R^2 (Within)	0.087	0.016

The table above plots results from regressions of the form

$$\log(Y_{i,t+k}) - \log(Y_{i,t}) = \alpha_0 + \alpha_t + \beta\eta_{i,t} + \epsilon_{i,t}$$

for $k = 20$ years for Census years spanning from 1850-2000. Here $Y_{i,t}$ is occupation's share in total non-farm employment. $\eta_{i,t}$ is standardized and growth rates are in annualized percentage terms. The sample is split into periods of innovation waves as identified by the breakthrough patent index of Kelly et al. (2020). The 20 year periods beginning in years 1880, 1910, 1920, and 1980, 1990 are labelled innovation waves with the remaining years representing non-innovation wave periods. Standard errors are clustered by occupation and corresponding t-stats are shown in parentheses. Observations are weighted by occupation employment share at time t.

Table 7: Technology And Employment At the Industry × Occupation Level (1910-2010)

	10 Years			20 Years		
$\eta_{i,t}$	-0.67***	-0.46***	-0.33**	-0.89***	-0.67***	-0.61***
	(-4.31)	(-3.45)	(-2.47)	(-5.45)	(-4.30)	(-3.70)
Time FE	X	X		X	X	
Industry FE		X			X	
Industry × Time FE			X			X
Emp Shares	X	X	X	X	X	X
N	101302	101302	101302	80027	80027	80025
R^2 (Within)	0.024	0.015	0.010	0.023	0.019	0.024

The table above plots results from regressions of the form

$$\log(Y_{i,j,t+k}) - \log(Y_{i,j,t}) = \alpha_0 + \beta\eta_{i,t} + \delta X_{t,j} + \epsilon_{i,j,t}$$

for $k = 10, 20$ years for Census years spanning from 1910-2010. Here $Y_{i,j,t}$ is the occupation i in industry j share of total non-farm employment. $\eta_{i,t}$ is standardized and growth rates are in annualized percentage terms. Standard errors are double clustered by occupation and industry, and corresponding t-stats are shown in parentheses. Observations are weighted by each occupation/industry pair's employment share at time t. The controls $X_{t,j}$ denote dummies for time, industry, or industry × time fixed effects as denoted by the specification, as well as the time t employment shares of the occupation-industry cell.

Table 8: Technology And Employment At the Industry × Occupation Level (Subsamples, 20-Year Horizon)

	1910-1950			1960-1990		
$\eta_{i,t}$	-0.86**	-0.74**	-0.68*	-0.84***	-0.50***	-0.53***
	(-2.51)	(-2.11)	(-1.94)	(-5.12)	(-3.04)	(-2.80)
Time FE	X	X		X	X	
Industry FE		X			X	
Industry × Time FE			X			X
Emp Shares	X	X	X	X	X	X
N	25260	25260	25258	54767	54767	54767
R^2 (Within)	0.027	0.035	0.053	0.026	0.0064	0.0073

The table above plots results from regressions of the form

$$\log(Y_{i,j,t+k}) - \log(Y_{i,j,t}) = \alpha_0 + \beta\eta_{i,t} + \delta X_{t,j} + \epsilon_{i,j,t}$$

for $k = 20$ years for the indicated subsamples. Here $Y_{i,j,t}$ is the occupation i in industry j share of total non-farm employment. $\eta_{i,t}$ is standardized and growth rates are in annualized percentage terms. Standard errors are double clustered by occupation and industry, and corresponding t-stats are shown in parentheses. Observations are weighted by each occupation/industry pair's employment share at time t. The controls $X_{t,j}$ denote dummies for time, industry, or industry × time fixed effects as denoted by the specification, as well as the time t employment shares of the occupation-industry cell.

Table 9: Regression of 1980-2012 Changes in Employment on $\eta_{i,1980}$ at Industry $\times$ Occupation Level

	(1)	(2)	(3)	(4)	(5)
$\eta_{i,1980}$	-0.76***	-0.80***	-0.72***	-0.72***	-0.76***
	(-5.32)	(-5.76)	(-5.01)	(-4.89)	(-5.51)
RTI		-0.23*			-0.19
		(-1.94)			(-1.48)
offshorability		-0.10			-0.14
		(-1.30)			(-1.62)
Robot Exposure			-0.58**		-0.68*
			(-2.00)		(-1.68)
Software Exposure				-0.28	0.043
				(-1.12)	(0.13)
Controls	X	X	X	X	X
Industry Fixed Effects	X	X	X	X	X

This table shows results from estimating

$$\Delta y_{i,j} = \alpha + \alpha_j + \beta\eta_{i,1980} + \delta X_i + \epsilon_{i,j}$$

Here i indexes occupations and j indexes industries. The outcome $\Delta y_{i,j}$ denotes the change in log employment over the 1980-2012 time period, reported in annualized percentage terms. Controls X_i include occupation employment share in 1980, occupation log wage in 1980, three categorical indicators for the occupation's average education level in 1980. We additionally include the routine-task intensity and the measure of occupation-level offshorability from Autor and Dorn (2013) and the measure of exposure to robots or software from Webb (2019) depending on the specification. Observations are weighted by employment share in 1980.

Table 10: Regression of 1980-2012 Changes in Wages on $\eta_{i,1980}$ at Industry $\times$ Occupation Level

	(1)	(2)	(3)	(4)	(5)
$\eta_{i,1980}$	-0.080***	-0.072***	-0.071***	-0.093***	-0.082***
	(-4.07)	(-3.52)	(-3.85)	(-4.92)	(-4.48)
RTI		-0.045			-0.035
		(-1.37)			(-1.07)
offshorability		0.0030			-0.0042
		(0.11)			(-0.15)
Robot Exposure			-0.12*		-0.27***
			(-1.82)		(-2.87)
Software Exposure				0.091	0.20***
				(1.62)	(4.70)
Controls	X	X	X	X	X
Industry Fixed Effects	X	X	X	X	X

This table shows results from estimating

$$\Delta y_{i,j} = \alpha + \alpha_j + \beta\eta_{i,1980} + \delta X_i + \epsilon_{i,j}$$

Here i indexes occupations and j indexes industries. The outcome $\Delta y_{i,j}$ denotes the change in log wages over the 1980-2012 time period reported in annualized percentage terms. Controls X_i include occupation employment share in 1980, occupation log wage in 1980, three categorical indicators for the occupation's average education level in 1980. We additionally include the routine-task intensity and the measure of occupation-level offshorability from Autor and Dorn (2013) and the measure of exposure to robots or software from Webb (2019) depending on the specification. Observations are weighted by employment share in 1980.

Table 11: Predicting 20-Year Growth Rates in Outcomes for NBER Manufacturing Industries Using $\psi_{j,t}$ (Index of Breakthrough Patents With High Avg Occupation Task Similarity)

Panel A: Industry Productivity and Output Measures

	TFP	Investment	Value Added	Value Added/Worker
$\psi_{j,t}$	1.48**	1.23**	1.06**	0.89***
	(2.23)	(2.41)	(2.57)	(3.36)
Controls	X	X	X	X
Year FE	X	X	X	X
N	1,930	1,930	1,930	1,930
R^2	0.63	0.40	0.39	0.41

Panel B: Industry Labor Share Measures

	Labor Share	Non-Prod Labor Share	Prod Labor Share	Prod Wage Share
$\psi_{j,t}$	-0.30*	-0.10	-0.58***	-0.29**
	(-1.82)	(-0.42)	(-3.49)	(-2.59)
Controls	X	X	X	X
Year FE	X	X	X	X
N	1,930	1,930	1,930	1,930
R^2	0.16	0.09	0.13	0.16

This table displays coefficients from regressions of the form

$$\log(X_{j,t+k}) - \log(X_{j,t}) = \alpha + \beta\psi_{j,t} + \delta Z_{j,t} + \epsilon_{j,t}$$

For $k = 20$ years and industry j for designated outcome variable X. Here $\psi_{j,t}$ is an index of industry breakthrough patents in year t and scaled by US population at time t. Breakthrough patents are assigned to industries using the patent CPC tech class to industry crosswalk from Goldschlag et al. (2020). The dependent variable is expressed in annualized percentage terms and $\psi_{j,t}$ is scaled to unit standard deviation. Controls $Z_{j,t}$ include industry employment shares, $\log(X_{j,t}) - \log(X_{j,t-5})$, $\log(X_{j,t})$, and year fixed effects. Standard errors are clustered by industry, and corresponding t-stats are shown in parentheses. The index $\psi_{j,t}$ only includes breakthrough patents that are above the median among all patent $\times$ industry pairs in the industry employment-weighted average of patent $\times$ occupation textual similarity. See main text for details.

5 Data Appendix

5.1 Converting Patent Text for Numerical Analysis

Here, we briefly overview our conversion of unstructured patent text data into a numerical format suitable for statistical analysis. We obtain text data for measuring patent/job task similarity from two sources. Job task descriptions come from the revised 4th edition of the Dictionary of Occupation Titles (DOT) database. We use the patent text data parsed from the USPTO patent search website in Kelly et al. (2020), which includes all US patents beginning in 1976, comprising patent numbers 3,930,271 through 9,113,586, as well as patent text data obtained from Google patents for pre-1976 patents. Our analysis of the patent text combines the claims, abstract, and description section into one patent-level corpus for each patent. Since the DOT has a very wide range of occupations (with over 13,000 specific occupation descriptions) we first crosswalk the DOT occupations to the considerably coarser and yet still detailed set of 6-digit occupations in the 2010 edition of O*NET. We then combine all tasks for a given occupation at the 2010 O*NET 6-digit level into one occupation-level corpus. The process for cleaning and preparing the text files for numerical representation follows the steps outlined below.

We first clean out all non-alphabetic characters from the patent and task text, including removing all punctuation and numerical characters. We then convert all text to lowercase. At this stage each patent and occupation-level task text are represented by a single string of words separated by spaces. To convert each patent/occupation into a list of associated words we apply a word tokenizer that separates the text into lists of word tokens which are identified by whitespace in between alphabetic characters. Since most words carry little semantic information, we filter the set of tokens by first removing all "stop words"– which include prepositions, pronouns, and other common words carrying little content–from the union of several frequently used stop words lists.

Stop words come from the following sources:

- `https://pypi.python.org/pypi/stop-words`
- `https://dev.mysql.com/doc/refman/5.1/en/fulltext-stopwords.html`
- `http://www.lextek.com/manuals/onix/stopwords1.html`
- `http://www.lextek.com/manuals/onix/stopwords2.html`
- `https://msdn.microsoft.com/zh-cn/library/bb164590`

- http://www.ranks.nl/stopwords
- http://www.text-analytics101.com/2014/10/all-about-stop-words-for-text-mining.html
- http://www.webconfs.com/stop-words.php
- http://www.nltk.org/book/ch02.html (NLTK stop words list)

We also add to the list of stop words the following terms that are ubiquitous in the patent text but don't provide information regarding the content and purpose of the patent: abstract, claim, claims, claimed, claiming, present, invention, united, states, patent, description, and background. The final stop word list contains 1337 unique terms that are filtered out.

Even after removing stop words, we expect much of the remaining text to offer little information regarding the purpose and use of a given patent or the core job functions expected to be performed by workers in a given occupation. In order to focus on the parts of the document most likely to contain relevant information, we retain descriptive and action words–i.e. nouns and verbs–and remove all other tokens. We do this using the part-of-speech tagger from the NLTK Python library. Finally, we lemmatize all remaining nouns and verbs, which is to convert them to a common root form. This converts all nouns to their singular form and verbs to their present tense. We use the NLTK WordNet Lemmatizer to accomplish this task. After these steps are completed, we have a set of cleaned lists of tokens for each patent and each occupation's tasks that we can then use to compute pairwise similarity scores.

5.2 Constructing a Statistical Displacement Factor

To construct our predictor we use a method proposed by Cong et al. (2019), which is well-suited to prediction exercises using large-scale textual data. Our adaptation of their method for the task of predicting occupation outcomes can be summarized in the following steps. Let the number of patent documents be N_p (where we restrict just to the set of breakthrough patents from Kelly et al. (2020) as described in section 1), the number of occupation task descriptions be N_o, and the number of words in the vocabulary formed from the union of all patent and occupation documents be N_w:

1. Perform approximate nearest neighbor search using a locality-sensitive hashing routine (LSH) on vector representations of word meanings to form K clusters ("topics") of related words. Label the kth cluster of words C_k.

2. Create a $N_p \times N_w$ matrix of breakthrough patent documents by word counts weighted by term-frequency inverse document frequency (TF-IDF), computed over all patents (i.e. TF-IDF is computed also including non-breakthrough patents). Call this matrix A. Loop over each word cluster C_k from step 1 for $k = 1, \ldots, K$, and extract the submatrix of A formed by taking the columns in A corresponding to the words contained in cluster C_k. Call this submatrix A_k. Perform a singular-value decomposition of A_k and take its top singular value v_k (in absolute value) and corresponding top right singular vector V_k. Then take the $N_p \times 1$ vector $P_k = \frac{|A_k v_k|}{v_k' v_k}$ to be the loadings of each patent document on topic/word cluster k. Retain only the clusters C_k which rank in the top 500 based on their top absolute singular values.

3. Perform step 2 for all occupations, except only for the top 500 clusters that were retained. Call the resulting $N_o \times 1$ vector of occupation loadings O_k. Denote the set breakthrough of patents issued in year t by $\widehat{\Gamma}_t$. Let $O_{k,i}$ represent the ith element of O_k and $P_{k,j}$ the jth element of P_k, the vector of patent loadings on cluster k. Then occupation i's exposure to the kth topic in year t is given by

$$\psi_{i,k,t} = \frac{O_{k,i}}{\kappa_t} \sum_{j \in \widehat{\Gamma}_t} P_{k,j} \tag{A.1}$$

As before we only sum over breakthrough patents and normalize by U.S. population in year t (denoted by κ_t). This yields an occupation's exposure in each year to the 500 topics which are found to be the most important among the breakthrough patents. Though equation A.1 looks a bit like our construction of $\eta_{i,t}$ in equation 12, it differs in that we no longer directly use word vectors to compute similarities. Instead, the Cong et al. (2019) technique only uses the word vectors to give an educated guess on the topics contained in the set of documents. Thus occcupations are similar to a given topic when they contain words that are also found in that topic.

We focus on the period of time covered by our CPS merged outgoing rotation group sample (1985-2018) used in the employment regressions in Figure 10. This is for two reasons: first, this is the period where our employment and wage data coverage is most comprehensive, with a yearly time series and relatively stable occupation classifications. Second, as documented in Figures 3, 4, and 5, the task composition of innovations has begun to change in this period of time relative to all previous innovation waves. In particular, cognitive skills have started to become more related to innovations, and this has been driven by the rising importance of information technology and electronics patents, which was not the case prior to the late 20th century. If skill-biased technological change has complemented the skillset of cognitive occupations, then innovations related to these occupations may be complementary to rather than a substitute for their skills. Thus if our measure mixes these two channels it is particularly likely to occur during this period of time.

Steps 1 and 2 above simply group documents into topics of related terms, compute how related a given topic is to each individual document, and provide an estimate of how important each topic is to the overall set of documents. Justification for the use of LSH clustering of word vectors to obtain topics and the singular value decomposition to infer topic importance/document topic loadings are discussed at length in Cong et al. (2019), to which we refer the interested reader for further details. For our purposes it suffices that by performing steps 1 through 3 we are able to obtain a panel of 500 predictors at the occupation-by-year level and which represent exposures to topics of words which are particularly relevant to patents.

Appendix Figures

Figure A.1: Sample Patent Topic Word Clusters

The above are four of the topics resulting from the LSH approximate nearest neighbors routine used to separate words into clusters as described in section 3.4. The relative size of the word corresponds to the importance of that word within the topic.

Appendix Tables

Table A.1: Predictive Performance of 10-Year Employment and Wage Growth on Predictors Constructed From Patent Topics

Panel A: Negative Constructed Predictor

	Employment Growth		Wage Growth	
ξ_{Mean}	-1.25***		-0.25***	
	(-6.68)		(-6.71)	
ξ_{PC1}		-1.10***		-0.24***
		(-6.32)		(-6.64)
Year FEs	X	X	X	X
Controls	X	X	X	X

Panel B: Positive Constructed Predictor

	Employment Growth		Wage Growth	
γ_{Mean}	0.87***		0.091***	
	(4.96)		(2.81)	
γ_{PC1}		0.64***		0.063***
		(3.93)		(2.80)
Year FEs	X	X	X	X
Controls	X	X	X	X

The tables above show coefficients from panel regressions of annualized wage and income growth rates over the 10-year horizon on textual factors constructed to predict employment as described in section 3.4. Regressions are of the form

$$y_{i,t+k} - y_{i,t} = \alpha + \beta Z_{i,t} + \delta X_{i,t} + \varepsilon_{i,t}$$

For $Z_{i,t} = \xi_{i,t}$ ("labor-saving") or $\gamma_{i,t}$ ("productivity enhancing"). Controls $X_{i,t}$ include three one-year lags of dependent variable, time fixed effects, wage, and occupation employment share. Subscripts $PC1$ and $Mean$ denote versions computed using either the first principal component or cross-sectional mean across individual textual predictors derived from the patent topics identified by the Cong et al. (2019) method. Dependent variable is expressed in annualized percentage terms and $\eta_{i,t}$ is standardized. Standard errors are clustered by occupation and independent variables are standardized. Observations are weighted by occupation's employment share at time t. The sample uses CPS merged outgoing rotation group data starting in 1982.

Table A.2: Correlations Between Predictors Constructed From Patent Topics and Different Versions of Occupation Technology Exposure $\eta_{i,t}$

	All Patents		Drop ICT Patents		Just ICT Patents	
ξ_{Mean}	0.74***		0.92***		0.47***	
	(22.47)		(34.91)		(12.39)	
γ_{Mean}		-0.0098		-0.12***		0.087*
		(-0.29)		(-5.11)		(1.84)

This table reports correlations between versions of technology exposure $\eta_{i,t}$ formed using different sets of patents and the composite predictors constructed from textual factors using the Cong et al. (2019) method to predict employment outcomes either negatively (ξ_{Mean}) or positively (γ_{Mean}). The "*Mean*" label denotes versions of composite predictors constructed by taking the cross-sectional means across individual textual factors which predict employment either negatively or positively. The first two columns represent the baseline measure of $\eta_{i,t}$ constructed using all patents; the next two columns drop ICT patents, defined to be those falling under the instruments/information or electronics categories; finally, the last two columns form $\eta_{i,t}$ only using ICT patents.